

# Breaking the glass ceiling: Configurational entropy measurements in extremely supercooled liquids

L. Berthier,<sup>1</sup> P. Charbonneau,<sup>2,3</sup> D. Coslovich,<sup>1</sup> A. Ninarello,<sup>1</sup> M. Ozawa,<sup>1</sup> and S. Yaida<sup>2</sup>

<sup>1</sup>*Laboratoire Charles Coulomb, CNRS-UMR 5221, Université de Montpellier, Montpellier, France*

<sup>2</sup>*Department of Chemistry, Duke University, Durham, North Carolina 27708, USA*

<sup>3</sup>*Department of Physics, Duke University, Durham, North Carolina 27708, USA*

Liquids relax extremely slowly on approaching the glass state. One explanation is that an entropy crisis, due to the rarefaction of available states, makes it increasingly arduous to reach equilibrium in that regime. Validating this scenario is challenging, because experiments offer limited resolution, while numerical studies lag more than eight orders of magnitude behind experimentally-relevant timescales. In this work we not only close the colossal gap between experiments and simulations but manage to create *in-silico* configurations that have no experimental analog yet. Deploying a range of computational tools, we obtain four estimates of their configurational entropy. These measurements consistently confirm that the steep entropy decrease observed in experiments is found also in simulations even beyond the experimental glass transition. Our numerical results thus open a new observational window into the physics of glasses and reinforce the relevance of an entropy crisis for understanding their formation.

*Introduction*—In his landmark 1948 paper, Kauzmann gathered experimental data for several glass-forming liquids and found that they all showed a steep decrease of their equilibrium configurational entropy upon lowering temperature towards their glass transition [1]. Theoretically, the nature of a thermodynamic glass transition associated with a vanishing configurational entropy is well-understood at the mean-field level [2, 3], suggesting that glass formation is accompanied by a rarefaction of available disordered states [4]. Its pertinence beyond the mean-field framework, however, remains controversial [5, 6]. In particular, it is still not known whether the entropy variation forms the core explanation of the glass transition. Experimental measurements are carried out over too limited a temperature range, within boundaries that have remained essentially unchanged since Kauzmann’s work and thus form a solid “glass ceiling.” In addition, experimental determinations of the configurational entropy are marred by approximations that influence their physical interpretation [7–9]. Computer simulations could potentially provide more precise estimates [10, 11], but have so far been restricted to a temperature range that is not experimentally relevant.

Can the debate over the role of configurational entropy ever be settled? At first sight, closure appears unlikely for two main reasons. (i) Measuring the configurational entropy below the experimental glass transition seems logically impossible, because experiments are constrained by their own duration, which fixes an upper limit to the accessible thermalization timescale,  $\tau$ . Specifically,  $\tau/\tau_0 \sim 10^{13}$  for molecules [12] (where the relaxation time at the onset temperature is  $\tau_0 \approx 10^{-10}$ s) and  $\tau/\tau_0 \sim 10^5$  for colloids [13] (where  $\tau_0 \approx 10^{-1}$ s). The situation for computer simulations is even worse. Current approaches access at most  $\tau/\tau_0 \sim 10^5$ , which is eight orders of magnitude behind molecular liquid experiments, and numerical progress has been slow. The two to three decades gained over the past 35 years [13–15] are mostly thanks to hardware improvements. At this pace, another century will be

needed before simulations attain experimentally-relevant conditions. The glass ceiling thus appears unbreakable. (ii) There is a fundamental methodological ambiguity as to which configurational entropy should be measured in order to match theoretical calculations. Qualitatively, the configurational entropy is defined by subtracting vibrational contributions from the total entropy [1, 10, 16]. What is specifically meant by “vibrations,” however, is ill-defined in general [4] and difficult to measure in practice [1, 11]. Hence consistently determining the configurational entropy is in itself a challenge.

Here, we solve these two major problems at once. First, we take advantage of the flexibility offered by computer simulations to dramatically accelerate the equilibrium sampling of configuration space [17, 18]. Namely, we use a system optimized for the nonlocal swap Monte Carlo (MC) algorithm, which enables its extremely fast thermalization. We establish that this approach surpasses any current alternative, and even experimental protocols. Second, we measure four proxies for the configurational entropy by deploying state-of-the-art computational tools to characterize *in-silico* configurations that are more deeply equilibrated than their experimental analogs [11, 18–20] and obtain consistent results with a clear physical interpretation. By combining these developments for a realistic model glass-former, we shift computer simulations from lagging eight decades behind experiments to exploring novel territory in glass physics. In particular, our measurements validate Kauzmann’s observations that the configurational entropy decreases steeply towards the glass temperature, and extend these observations to a regime previously inaccessible.

*Results*—We simulate a three-dimensional polydisperse mixture of hard spheres, as in [18], which is a good model for colloids used in experiments [13, 21]. We show in Appendix H that our methods and conclusions also apply to particles with soft and more complex interactions. We control the volume fraction  $\phi$ , and measure pressure  $P$  to report the (unitless) reduced pressure,  $Z = P/(\rho k_B T)$ ,

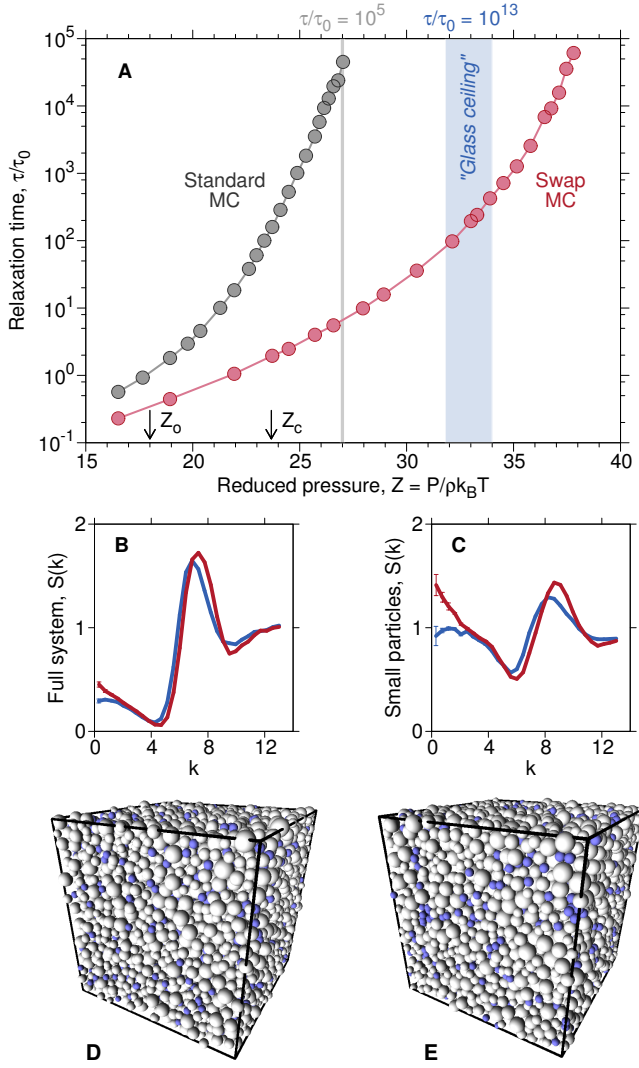


FIG. 1. Breaking the glass ceiling: thermalization beyond the experimental glass transition. (A) Structural relaxation time  $\tau$  for both standard and accelerated swap MC dynamics as a function of the reduced pressure of a polydisperse hard sphere model.  $\tau$  is obtained from the decay of the self-intermediate scattering function at a wavenumber  $k = 5.25$ . The onset of slow dynamics is at  $Z_0 \approx 18$ , and the mode-coupling crossover at  $Z_c \approx 23.5$ . Times are rescaled by  $\tau_0 = \tau(Z_0)$  for standard MC. The current limit of colloidal ( $\tau/\tau_0 = 10^5$ ) and molecular ( $\tau/\tau_0 = 10^{13}$ ) experiments are indicated by vertical bands (this uncertainty stems from the extrapolation scheme), showing that swap MC breaks the glass ceiling. Static structure factor for  $Z = 18.8 \approx Z_0$  ( $\phi = 0.568$ ) and  $Z = 33.2$  ( $\phi = 0.640$ ) for (B) all particles (C) the 40% particles with the smallest diameter. (D, E) show typical snapshots for these two state points, where the smallest particles are highlighted in blue.

where  $\rho$  is the number density, and  $k_B T$  the thermal energy. This natural control variable for hard spheres plays a role akin to the inverse temperature in thermal liquids [22]. Detailed information about the simulations

is provided in Appendices A, B, and C. Swap MC complements standard translational MC moves with nonlocal moves that exchange randomly-chosen pairs of particles, ensuring equilibrium sampling. We demonstrate the extreme speedup actually achieved by swap MC for this model in Fig. 1, in which the structural relaxation time  $\tau$  for both MC sampling methods is reported as the system approaches its glass transition. Note that the rapid increase of  $\tau$  in standard MC simulations resembles the fragile super-Arrhenius behavior of standard glass-formers [4]. We have thus fitted several empirical forms to our measurements, which thermalize up to  $Z \approx 27$ , to estimate the experimental glass transition at  $\tau/\tau_0 = 10^{13}$  (see Appendix D). The fits give consistent locations for the glass ceiling,  $Z_g \approx 32$ -34, as highlighted in Fig. 1. Remarkably, this dramatic slowdown is completely bypassed by swap MC sampling, which thermalizes the system up to  $Z \approx 38 > Z_g$ . Even our most conservative extrapolation indicates that we access a dynamical range that is broader than in experiments. Meanwhile, the two-point structure barely budes (see Figs. 1B-C), which is a tell-tale sign of glassiness [4] and a confirmation that both crystallization and more subtle fractionation effects are absent. Visual inspection of particle configurations further confirm these conclusions, see *e.g.* Figs. 1D-E. We are therefore in the unique position of studying at equilibrium a homogeneous fluid beyond the experimental glass ceiling.

We then turn to the measurement of the configurational entropy,  $s_{\text{conf}}$ , in these extremely supercooled configurations. The numerical procedures leading to the four estimates of  $s_{\text{conf}}$  are shown in Fig. 2. Further details are provided in Appendices E, F and G. In Method 1, we determine the configurational entropy from its most conventional definition,  $s_{\text{conf}} = s_{\text{tot}} - s_{\text{vib}}$ , as used in many experimental and simulation studies [1, 10, 16, 19]. The total entropy of the equilibrium fluid,  $s_{\text{tot}}$ , is measured by thermodynamic integration from the dilute ideal gas limit to the target volume fraction, while the vibrational contribution,  $s_{\text{vib}}$ , is measured by Frenkel-Ladd thermodynamic integration [19, 23]. The latter integration is over the amplitude of the Hookean constant,  $\alpha$ , of a spring that constrains each particle to reside close to the position of a quenched reference equilibrium configuration. This requires estimating the mean-squared distance  $\delta r^2$  between the reference and constrained systems over a broad range of  $\alpha$  values, as illustrated in Fig. 2A. In continuous polydisperse systems, special care is also needed to account for the mixing contribution to the total entropy, because it formally diverges [24, 25]. We measure the mixing entropy from an independent, additional set of simulations [25].

Methods 2 and 3 are both based on the Franz-Parisi theoretical construction [26], which expresses the equilibrium free energy of the liquid,  $V(Q)$ , in terms of a global order parameter, the overlap  $Q$ . The overlap between two configurations is defined as  $Q = N^{-1} \sum_{i,j} \theta(a - |\mathbf{r}_{1,i} - \mathbf{r}_{2,j}|)$ , where  $\theta(x)$  is the Heaviside function,  $\mathbf{r}_{1,i}$

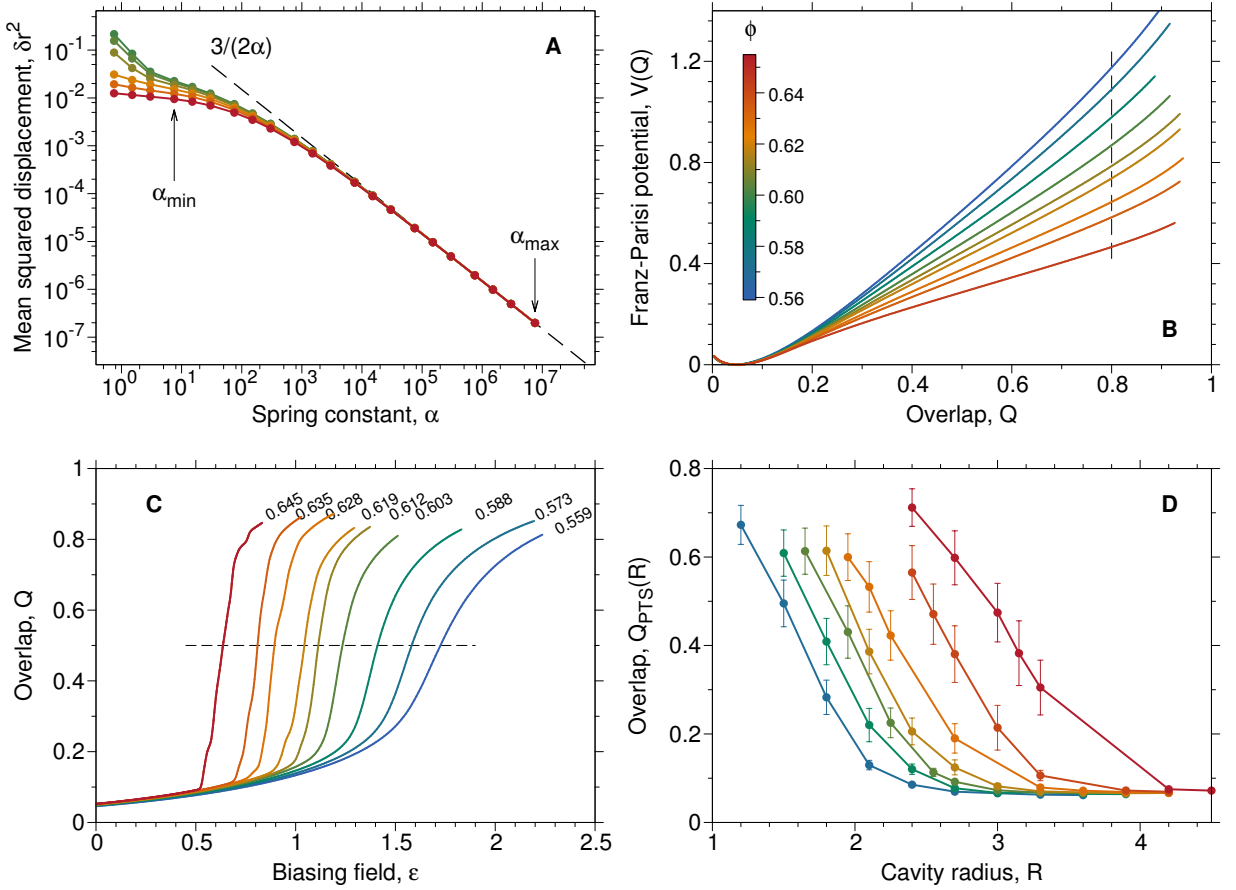


FIG. 2. Numerical procedures leading to the four estimates of the configurational entropy. (A) Method 1: The Frenkel-Ladd method to obtain the vibrational entropy  $s_{\text{vib}}$  performs a thermodynamic integration of the mean-squared distance  $\delta r^2$  between a reference equilibrium configuration and a copy of the system constrained by a harmonic potential of strength  $\alpha$ . The integration is carried out from  $\alpha_{\max}$ , for which the system behaves as an Einstein solid (indicated by the dashed line  $\delta r^2 = 3/(2\alpha)$ ) to  $\alpha_{\min}$ , for which particles are trapped by their own cages on the vibrational time scale. (B) Method 2: The numerically-determined Franz-Parisi potential  $V(Q)$  is used to measure the configurational entropy as  $s_{\text{conf}} = V(Q_{\text{high}} = 0.8) - V(Q_{\text{low}} \approx 0.05)$ . (C) Method 3: The evolution of the overlap  $Q$  with the biasing field  $\varepsilon$  reveals a first-order jump at a value  $\varepsilon^*$  for which  $Q = 1/2$  (dashed line). Then,  $s_{\text{conf}} = \varepsilon^*(Q_{\text{high}} - Q_{\text{low}})$ . (D) Method 4: The decay of the cavity overlap correlation function  $Q_{\text{PTS}}(R)$  with cavity radius,  $R$ , defines the point-to-set correlation length  $\xi_{\text{PTS}} \propto 1/s_{\text{conf}}$ .

and  $\mathbf{r}_{2,j}$  are the positions of particle  $i$  and  $j$  within configuration 1 and 2, and  $a$  is a fraction of the average particle diameter. By definition,  $Q$  quantifies the similarity between the density profiles of two configurations. To compute  $V(Q)$ , we introduce a coupling between a quenched reference equilibrium configuration and a copy of the system through a field  $\varepsilon$  conjugate to  $Q$  [11, 26];  $\varepsilon$  constrains the collective density profile, whereas  $\alpha$  in Method 1 constrains single-particle displacements. We define  $V(Q) = -\lim_{\varepsilon \rightarrow 0} \left[ \frac{T}{N} \ln P(Q) \right]$ , where  $P(Q)$  is the equilibrium probability distribution of the overlap for a given reference configuration, and brackets denote averaging over these configurations. In Method 2, we follow [11] and use the free-energy difference  $s_{\text{conf}} = V(Q_{\text{high}}) - V(Q_{\text{low}})$  between the global minimum at  $Q_{\text{low}} \approx 0.05$  and its value at  $Q_{\text{high}} = 0.8$  to obtain an estimate of  $s_{\text{conf}}$  that is closest to its theoretical definition, see Fig. 2B. Importantly, this estimate only

exists for sufficiently supercooled states, for which  $Q_{\text{high}}$  can be defined [11]. For the present system this happens close to the mode-coupling crossover,  $Z_c$ . In Method 3, we determine the value of the biasing  $\varepsilon$  needed to ‘tilt’ the potential  $V(Q)$ , so that a first-order phase transition occurs where  $Q$  jumps from  $Q_{\text{low}}$  to  $Q_{\text{high}}$ , as illustrated in Fig. 2C. We use the maximum variance of the overlap fluctuations to measure  $\varepsilon^*$  for each volume fraction studied. In practice, this is equivalent to determining the biasing field at which the overlap reaches  $Q = 1/2$ , see Fig. 2C.

Method 4 builds on the physical idea that the decrease of the configurational entropy is directly responsible for the growth of spatial correlations quantified by the point-to-set correlation length,  $\xi_{\text{PTS}}$  [27–29]. Following what is now common practice [20, 29], we measure  $\xi_{\text{PTS}}$  by pinning the position of particles outside a spherical cavity of radius  $R$ , equilibrating the liquid within it, and measur-

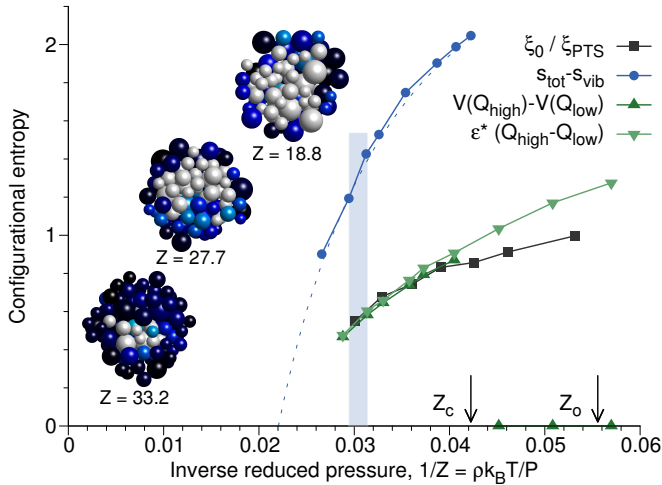


FIG. 3. Convergent measurements of the four estimates of the configurational entropy beyond the glass ceiling.  $s_{\text{conf}}$  is plotted as a function of  $1/Z \propto T/P$ , which is equivalent to the classic Kauzmann plot. All measurements indicate a steep decrease of  $s_{\text{conf}}$  which continues as the experimental glass ceiling is crossed. The dashed line is an extrapolation based on  $s_{\text{tot}} - s_{\text{vib}}$  (see Appendix E). Inset: typical overlap profiles measured in a finite cavity of radius  $R = 3.0$ , with colors coding for the overlap value from low (white) to large (black). Overlap fluctuations are uncorrelated around the onset but become strongly correlated over the entire cavity at the largest shown pressure.

ing the evolution of the overlap between interior configurations,  $Q_{\text{PTS}}(R)$ , with the cavity radius  $R$ , as shown in Fig. 2D. The decay of  $Q_{\text{PTS}}(R)$  is controlled by  $\xi_{\text{PTS}}$ , and the variance of the overlap fluctuations also presents a maximum [20] very close to  $\xi_{\text{PTS}}$ . Physically,  $\xi_{\text{PTS}}$  thus represents the cavity size above which the system starts to explore a significant number of distinct states. With minimal hypothesis [28], it can be connected to the configurational entropy through  $s_{\text{conf}} \propto \xi_{\text{PTS}}^{-(d-\theta)}$ , which involves an unknown exponent  $\theta \leq (d-1)$ . We make the simplest assumption that the inequality is saturated, and define the configurational entropy as  $s_{\text{conf}} \propto 1/\xi_{\text{PTS}}$ . The value  $\xi_0 = 1.69$  of  $\xi_{\text{PTS}}$  at  $Z = Z_0$  fixes the prefactor.

We gather the four estimates of the configurational entropy in Fig. 3 to produce a plot equivalent to the original 1948 Kauzmann representation of  $s_{\text{conf}}(T)$  [1]. In the high-temperature liquid, the configurational entropy is not sensibly defined [11], although three of the four measures can still be estimated there. Only this regime was studied in earlier simulations [10, 11, 19]. In the more interesting low-temperature regime, our main finding is that the important conceptual and technical differences between the four methods nevertheless result in qualitatively consistent results. In particular, the three estimates (Methods 2-4) that closely follow the theoretical definition of the configurational entropy provide numerically indistinguishable results at low temperatures. The conventional estimate of the entropy (Method 1) is

larger, as expected [30], but its temperature evolution remains qualitatively consistent with the other methods. All data estimates of  $s_{\text{conf}}$  thus exhibit a steep decrease as  $Z$  increases towards the glass phase. Although quantitative extrapolation is hard to control, our measurements robustly suggest that  $s_{\text{conf}}$  may vanish near  $Z \approx 1/0.022 \approx 45$ . We thus conclude that even for a simple glass-forming system equilibrated deeper in the landscape than any previously studied material, the trend discovered 70 years ago by Kauzmann is confirmed when more precise estimates of  $s_{\text{conf}}$  are adopted, and persists even below the experimental glass temperature. We further show in Appendix H that similar observations can be performed for a model with a continuous pair potential, suggesting our methodological progress and physical conclusions are not restricted to hard spheres.

*Discussion*—Our point-to-set measurements go beyond Kauzmann’s observation as they further establish that the decrease in  $s_{\text{conf}}$  is accompanied by an increase of static spatial correlations as the glass ceiling is crossed, a result which reinforces a recent experimental report based on non-linear dielectric measurements [31]. Our particle-based resolution of such correlations further provides a direct visualization of the spatial profile of the overlap within a spherical cavity (see in Fig. 3 insets). In particular, within a cavity comprising about 200 particles, positions of particles freely fluctuate near the onset pressure, whereas they become strongly correlated over the entire cavity for the largest pressure shown. The spatial extent of static correlations is thus directly revealed.

The important methodological progress achieved here regarding the thermalization of supercooled liquids therefore support a thermodynamic view of the glass formation based on the rarefaction of metastable accompanied by growing static correlations that is more direct than any work since Kauzmann’s.

## ACKNOWLEDGMENTS

The research in Montpellier was supported by funding from the European Research Council under the European Unions Seventh Framework Programme (FP7/2007-2013) / ERC Grant agreement No 306845 and by a grant from the Simons Foundation (#454933, Ludovic Berthier). The research at Duke was supported by a grant from the Simons Foundation (#454937, Patrick Charbonneau).

## Appendix A: Model

We study a three-dimensional hard-sphere model, for which the pair interaction is zero for non-overlapping particles and infinite otherwise. Systems have a continuous size polydispersity, with particle diameters  $\sigma$  randomly drawn from the distribution  $f(\sigma) = A\sigma^{-3}$ , with  $\sigma \in [\sigma_{\min}, \sigma_{\max}]$  and normalization constant  $A$ . Our model is the same as in Ref. [18], with a measure of size polydispersity  $\Delta = \sqrt{\overline{\sigma^2} - \bar{\sigma}^2}/\bar{\sigma}$ , where  $\overline{\cdots} = \int d\sigma f(\sigma)(\cdots)$ , of  $\Delta = 23\%$ , and  $\sigma_{\min}/\sigma_{\max} = 0.4492$ . The average diameter,  $\bar{\sigma}$ , defines the unit of length. We simulate systems composed of  $N$  particles in a cubic cell of volume  $V$  under periodic boundary conditions [32]. Depending on the chosen method to estimate the configurational entropy we simulate either  $N = 1000, 8000$  (Method 1) or  $N = 300$  (Method 2 and 3), and cavities for Method 4 are carved from bulk configurations with  $N = 8000$ . The relaxation times shown in Figure 1A are obtained from samples with  $N = 1000$ . Given these parameters, the system is then uniquely characterized by its volume fraction  $\phi = \pi N \bar{\sigma}^3/(6V)$ , and we frequently report the data using the reduced pressure  $Z = P/(\rho k_B T)$ , where  $\rho$ ,  $k_B$ , and  $T$  are the number density, Boltzmann constant and temperature, respectively. Without loss of generality, we set  $k_B$  and  $T = 1/\beta$  to unity. The pressure  $P$  is calculated from the contact value of the pair correlation function properly scaled for a polydisperse system [33].

## Appendix B: Methods

To obtain equilibrium fluid configurations deep in the glassy regime, we perform Monte-Carlo (MC) simulations with both translational displacements and non-local particle swaps [17, 18, 34–40]. The two types of moves are selected randomly: with probability 0.8 we attempt a translational displacement, and with probability 0.2 we attempt a swap. Translational displacements are uniformly drawn over a cube of side 0.115. For swaps, two randomly selected particles exchange diameter. In both cases, proposed moves are accepted if no overlap is created.

We measure the equilibrium relaxation time  $\tau$  both with and without the swap moves from the time-decay of the self-intermediate scattering function,  $F_s(k, \tau) = 1/e$ , where  $k = 5.25$  is the wavenumber chosen slightly below the first maximum of the static structure factor. Note that the particles' diameters can change during a swap MC simulation, but their trajectories are continuous. Relaxation times are measured in units of MC sweeps, comprising  $N$  MC moves, irrespective of their type.

Thermalized systems at each state point are obtained in the same way for both standard and swap MC dynamics. We measure the relaxation time  $\tau$  and ensure that for each state point simulations of a total duration of at least  $100\tau$  can be performed. We also check for the presence of aging effects in time correlation functions, and we measure the static structure factor, the pair correlation function and the equation of state over long simulations, paying attention to any temporal drift that could signal either improper thermalization, incipient crystallization or demixing of particles with distinct sizes. Selected results for the evolution of the structure factor with volume fraction are presented in Figs. 1B-C. Over the extreme range of densities shown here, the static structure evolves very little. Similarly, a very modest evolution is seen when the partial structure factor of the smallest particles is measured. A large increase of the low- $k$  value of these quantities, or the emergence of discrete peaks would signal that demixing or crystallization is taking place. In fact, we have found that measuring the relaxation time for the swap simulation is the most sensitive test of thermalization, because purely static observables may appear thermalized over long simulation times, whereas the system is in fact nearly arrested in a glass state within which sampling is inefficient.

We introduce two dynamical reference states: (i) the onset of slow dynamics at  $\phi_0 \approx 0.56$  ( $Z_0 \approx 18$ ), above which the time decay of correlation functions is non-exponential [41], and (ii) the mode-coupling crossover  $\phi_c \approx 0.598$  ( $Z_c \approx 23.5$ ), at which a power-law fit extrapolates a divergence of the relaxation times [18]. Note that these particular definitions are not unique [42], but are sufficiently accurate for our purposes, where these values are simply used for qualitative reference.

## Appendix C: Equation of state

For reference, we report in Fig. S1 the equilibrium equation of state (from Ref. [18]) of the system under study,  $Z = Z(\phi)$ , thus enabling the translation of results from  $Z$  to  $\phi$ . Also, the data points used in four measurements of the configurational entropy are presented in Tab. I, II, and III.

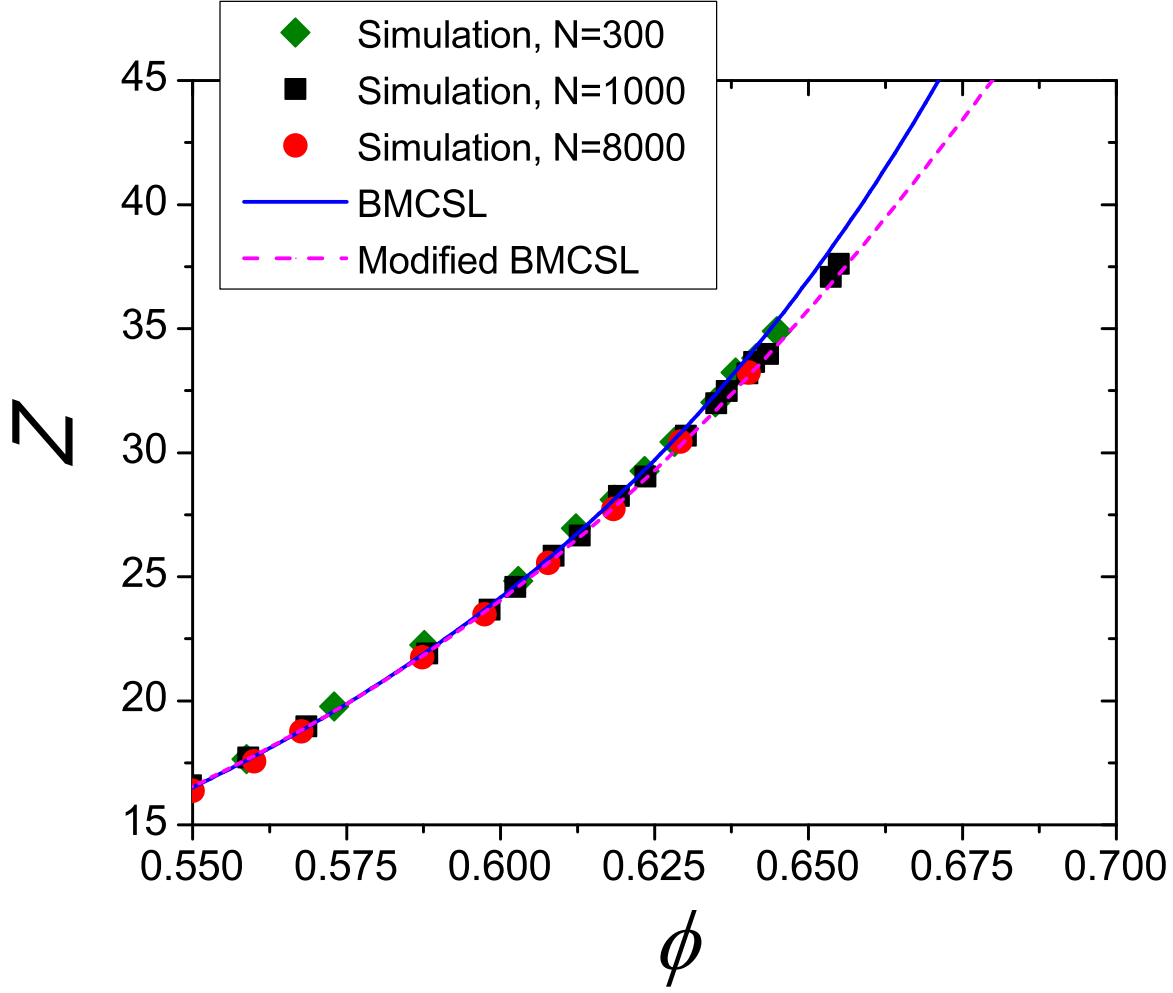


FIG. S1. Equation of state  $Z = Z(\phi)$  for the studied hard-sphere system. The solid line is the empirical expression for the equation of state of Refs. [43, 44], BMCSL, which describes our data well, except at very large volume fractions where it slightly overestimates  $Z$ . A modified version of BMCSL [45] is also shown by the dashed line.

$Z$	17.6	19.8	22.3	24.8	27.0	28.1	30.4	32.0	34.9
$\phi$	0.559	0.573	0.588	0.603	0.612	0.619	0.628	0.635	0.645

TABLE I. Data points for  $N = 300$  used in the Franz-Parisi construction (Methods 2 and 3).

$Z$	23.7	24.6	25.8	28.3	30.7	32.0	34.0	37.6
$\phi$	0.598	0.602	0.609	0.619	0.630	0.635	0.643	0.655

TABLE II. Data points for  $N = 1000$  used in the thermodynamic integrations (Methods 1).

$Z$	18.8	21.7	23.5	25.6	27.7	30.4	33.2
$\phi$	0.568	0.587	0.597	0.608	0.618	0.629	0.640

TABLE III. Data points for  $N = 8000$  used in the thermodynamic integrations (Methods 1) and the point-to-set correlation (Method 4).



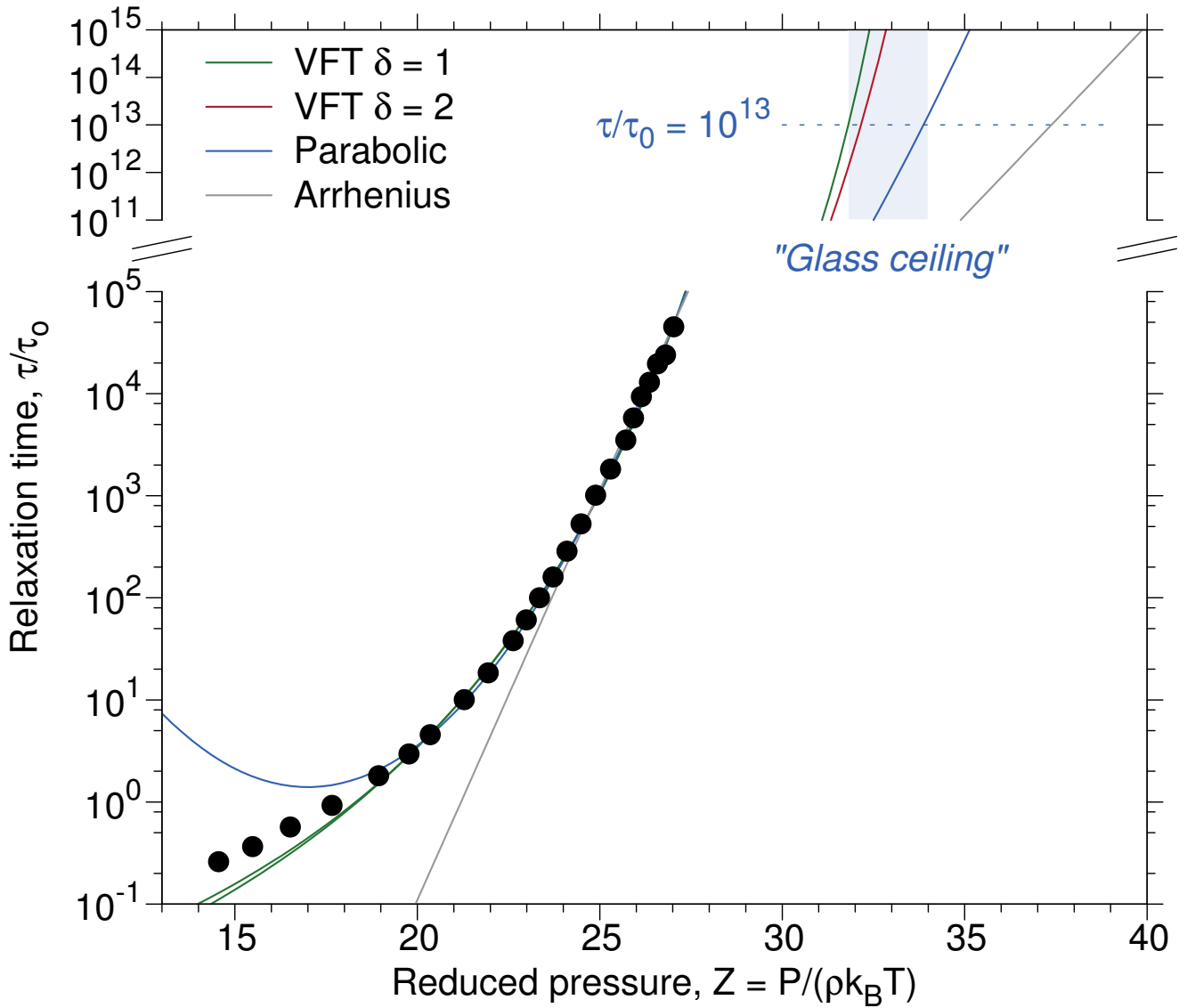


FIG. S2. The measured relaxation time,  $\tau$ , (symbols) is fitted to the VFT form, Eq. (D1), with exponents  $\delta = 1$  and  $\delta = 2$ , to the singularity free parabolic law, Eq. (D2), and to the Arrhenius law, Eq. (D3). The first three fits provide a good description of the data, and are used to locate the glass ceiling at  $\tau/\tau_0 = 10^{13}$  that is used in the main text.

#### Appendix D: Locating the glass ceiling

In the context of this work, we refer to the typical relaxation time measured at the laboratory glass transition (at which, conventionally,  $\tau/\tau_0 = 10^{13}$  [12]) as the glass ceiling,  $Z_g$ . Because standard MC dynamics can only access relaxation times at most of order  $\tau/\tau_0 = 10^5$ , where  $\tau_0 = 10^4$  MC sweeps is the value of  $\tau$  at the onset of slow dynamics, our dynamic data has to be extrapolated to locate this ceiling. We fit measured relaxation times to various functional forms up to  $\tau/\tau_0 \leq 10^5$ , and then extrapolate up to the vicinity of the glass ceiling. In an effort to obtain an estimate as unbiased as possible, we consider a range of possible functional forms [4, 22], as detailed below. We present in Fig. S2 the result of this exercise. In particular, note that whereas extrapolating the location of a putative divergence of  $\tau$  is extremely delicate, extrapolating the location of the glass ceiling is in fact rather well-controlled.

The first functional form is the Vogel-Fulcher-Tammann (VFT) expression

$$\tau = \tau_\infty \exp \left[ \frac{A}{(Z_{\text{vft}} - Z)^\delta} \right], \quad (\text{D1})$$

where  $\tau_\infty$ ,  $A$ , the exponent  $\delta$  and the critical pressure  $Z_{\text{vft}}$  are free parameters. Whereas  $\delta = 1$  is traditionally used,

more recent experimental and numerical studies [13, 22] favor  $\delta = 2$ .

In Fig. S2 we consider both. As noted before [18], the fit with  $\delta = 1$  yields a critical pressure  $Z_{\text{vft}} \approx 38$  that falls within the range in which swap MC sampling equilibrates. Although the resulting fit is good over the range covered by standard MC dynamics, it overestimates the growth of the relaxation time beyond that range, which undermines the very logic behind the proposed scaling form. When  $\delta = 2$  is imposed the fit is still very good, but we now obtain  $Z_{\text{vft}} \approx 45$ , which is at least beyond the equilibrium range accessible with the swap method. Although this fit provides a more consistent description of the dynamical data, resolving one form from the other is beyond the scope of the current study.

The second functional form is the parabolic law proposed by Elmatad *et al.* [46, 47] in the context of facilitated models,

$$\tau = \tau_{\infty} \exp [A(Z - Z_0)^2], \quad (\text{D2})$$

where  $Z_0 = 17$  is around the onset of slow dynamics and the fit is made over the range  $Z > Z_0$ . In contrast to the VFT law, this parabolic expression does not invoke a divergence of the relaxation time at any finite pressure;  $\tau$  only diverges when  $Z$  also diverges. Note that because  $Z \propto 1/T$ , this form is equivalent to fitting the relaxation time of a supercooled liquid without any finite temperature divergence [22]. As a result this expression necessarily provides a less divergent extrapolation of the relaxation time at large pressures, but still accounts well for the curvature, *i.e.*, the fragility, of the measured dynamical data. The leading order at large  $Z$ ,  $\tau \propto \exp[AZ^2]$ , indeed grows faster than the Arrhenius law. Like the VFT form, this fit is quite good over the measured range of relaxation times, see Fig. S2.

For completeness, we also include a simple Arrhenius fit to the high- $Z$  portion of the equilibrium data

$$\tau = \tau_{\infty} \exp[AZ]. \quad (\text{D3})$$

We find this fit not to be very good as it does not capture the fragility of the system. It is also inconsistent with the thermodynamic behavior of the system reported in the main text, because the configurational entropy typically does not vary much in glass-formers with an Arrhenius-like behavior. This fit is thus most probably incorrect in the sense that it underestimates considerably the evolution of the relaxation time of the system.

We conclude that the first two families of expressions account well for the non-Arrhenius dependence of the relaxation time data observed in Fig. S2 and yield reasonable descriptions of the available equilibrium data over several orders of magnitude. Whereas the VFT law with  $\delta = 1$  can be logically ruled out by the swap MC measurements, the other two fits cannot be excluded on the bases of any further measurement we could perform. As is common in glass simulations, it is thus difficult to discriminate between fits with or without a finite temperature singularity [4]. These forms nonetheless allow us to locate the glass ceiling  $Z_g$  for which the relaxation time is  $10^{13}$  larger than its value at  $\phi_0$ . Fits to the relevant expressions yield estimates for the glass ceiling ranging from  $Z_g \approx 32$  (VFT law with  $\delta = 1$ ) to  $Z_g = 34$  (parabolic law). These values are used to draw the glass ceiling box in Fig. 1 of the main text.

As discussed above, there are several reasons for which we do not expect the Arrhenius expression to describe accurately the relaxation data at high  $Z$ . If we nonetheless considered it, the equilibrium swap MC data would still be able to equilibrate the liquid at pressures higher than the glass transition it predicts. Hence, this very conservative extrapolation allows us to confidently state that we have successfully broken the glass ceiling.

## Appendix E: Method 1—Conventional definition

In order to compute the configurational entropy,  $s_{\text{conf}}$ , we define

$$s_{\text{conf}} = s_{\text{tot}} - s_{\text{vib}} \quad (\text{Method 1}), \quad (\text{E1})$$

where  $s_{\text{tot}}$  and  $s_{\text{vib}}$  are the total and vibrational entropy, respectively. Note that this definition is common in experimental and computational studies [1, 10, 16, 19, 48, 49].

### 1. Total entropy

The total entropy is obtained by thermodynamic integration from the ideal gas limit ( $\phi \rightarrow 0$ ) up to the target volume fraction  $\phi$ ,

$$s_{\text{tot}}(\phi) = \frac{5}{2} - \ln \left( \frac{6\phi}{\pi M_3} \right) - \ln \Lambda^3 - \int_0^\phi d\phi' \frac{(Z(\phi') - 1)}{\phi'} + s_{\text{mix}}, \quad (\text{E2})$$



where  $M_k = \overline{\sigma^k}$  ( $k = 1, 2, 3, \dots$ ) are the  $k$ -th moments and  $\Lambda$  is the thermal de Broglie wavelength. Without loss of generality, we here set  $\Lambda = 1$ . For continuous polydisperse systems one needs to pay special attention to the mixing entropy,  $s_{\text{mix}}$  [24, 50]. We get back to this point in Sec. E3. For now, we report  $s_{\text{tot}} - s_{\text{mix}}$  for  $N = 1000$  and  $8000$  in Fig. S3(A).

In order to validate the numerical thermodynamic integration, we also consider an analytical approximation of the equation of state (EOS). The polydisperse version of the Carnahan-Starling EOS, *i.e.*, the so-called BMCSL EOS [43, 44],

$$Z_{\text{BMCSL}}(\phi) = \frac{1}{1-\phi} + \frac{3M_1M_2}{M_3} \frac{\phi}{(1-\phi)^2} + \frac{M_2^3}{M_3^2} \frac{(3-\phi)\phi^2}{(1-\phi)^3}, \quad (\text{E3})$$

is known to describe experiments and simulations of polydisperse hard-sphere systems rather well in the liquid regime. We also consider a modified version of  $Z_{\text{BMCSL}}$ ,

$$Z_{\text{modBMCSL}}(\phi) = 1 + h(\phi)(Z_{\text{BMCSL}}(\phi) - 1), \quad (\text{E4})$$

where  $h(\phi) = 0.005 - \tanh(14(\phi - 0.79))$  [45]. Both EOSs trace our simulation data very well, as can be seen in figs. S1 and S3(A). These EOSs can thus also be used to extrapolate the configurational entropy toward very high volume fraction (Sec. E4).

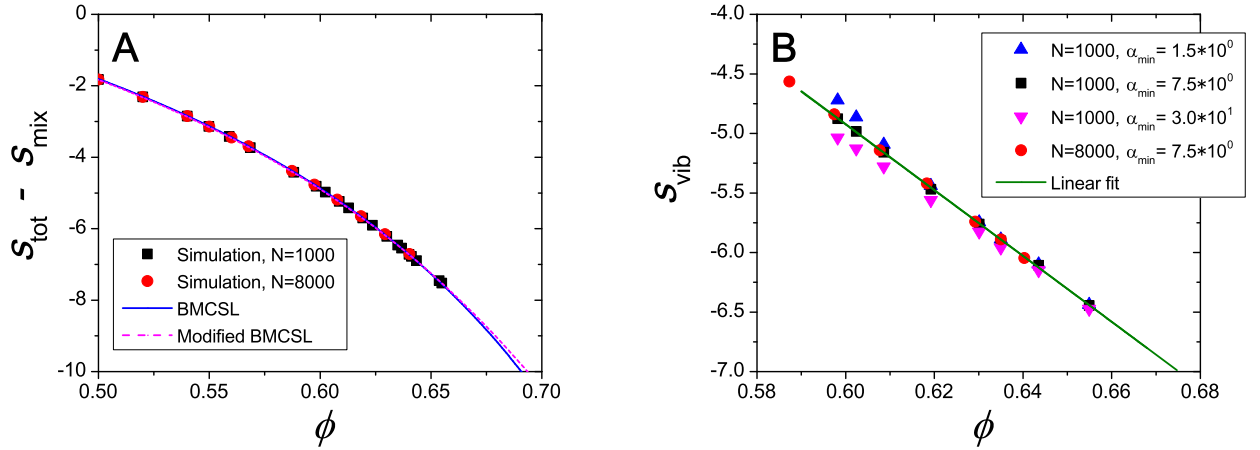


FIG. S3. (A) The total entropy minus the mixing contribution from simulations and from the (modified) BMCSL EOS. (B)  $s_{\text{vib}}$  as a function of  $\phi$  for several  $\alpha_{\text{min}}$ . The solid straight line is fitted to the  $N = 1000$  results with  $\alpha_{\text{min}} = 7.5 \times 10^0$  in the high  $\phi$  region.

## 2. Vibrational entropy

The vibrational entropy,  $s_{\text{vib}}$ , is obtained by Frenkel-Ladd (FL) thermodynamic integration [19, 38] by performing MC simulations of a constrained system with Hamiltonian

$$\beta H(\alpha) = \beta H(0) + \alpha \sum_{i=1}^N (\mathbf{r}_i - \mathbf{r}_{0i})^2, \quad (\text{E5})$$

for a template configuration,  $\{\mathbf{r}_{0i}\}$ , obtained from an equilibrium target system under  $H(0)$ . In short, the FL method integrates from a large  $\alpha_{\text{max}} \gg 1$ , at which particles experience a nearly pure harmonic oscillator, down to a very weak  $\alpha_{\text{min}} \ll 1$ , at which particles vibrate within cages. The vibrational entropy is then obtained by

$$s_{\text{vib}} = \frac{3}{2} - \ln \Lambda^3 - \frac{3}{2} \ln \left( \frac{\alpha_{\text{max}}}{\pi} \right) + \int_{\alpha_{\text{min}}}^{\alpha_{\text{max}}} d\alpha \delta r^2(\alpha) + \alpha_{\text{min}} \delta r^2(\alpha_{\text{min}}), \quad (\text{E6})$$

$$\delta r^2(\alpha) = \left[ \left\langle \frac{1}{N} \sum_{i=1}^N (\mathbf{r}_i - \mathbf{r}_{0i})^2 \right\rangle_{\alpha} \right], \quad (\text{E7})$$

where  $\langle(\dots)\rangle_\alpha$  and  $[(\dots)]$  denote the thermal average with Hamiltonian  $H(\alpha)$  and averaging over template configurations,  $\{\mathbf{r}_{0i}\}$ , respectively.  $\delta r^2$  in Eq. (E7) is the mean-squared displacement shown in Fig. 2(A) of the main text.

Numerical integration of  $\delta r^2$  is performed from  $\alpha_{\max}$  to  $\alpha_{\min}$ . We set  $\alpha_{\max} = 7.5 \times 10^6$ , which is well into the harmonic-oscillator scaling regime [see Fig. 2(A) of the main text],

$$\delta r^2(\alpha_{\max}) = \frac{3}{2\alpha_{\max}}. \quad (\text{E8})$$

We choose  $\alpha_{\min}$  such that  $\delta r^2$  of the constrained system is comparable to the mean-squared displacement of the target system without constraint. The last term of Eq. (E6) is included, assuming a constant  $\delta r^2$  below  $\alpha_{\min}$  and using zero as the lower limit of the integral in Eq. (E6). In Fig. S3(B), we show  $s_{\text{vib}}$  for  $N = 1000$  and  $8000$ , with  $\alpha_{\min} = 1.5 \times 10^0$ ,  $7.5 \times 10^0$ , and  $3.0 \times 10^1$ . At high volume fraction  $\phi$ , the results are insensitive to the choice of  $\alpha_{\min}$ , as expected. We can thus confidently set  $\alpha_{\min} = 7.5 \times 10^0$ . Also, we empirically observe a linear relation between  $s_{\text{vib}}$  and  $\phi$ , which allows us to linearly extrapolate the fit. We use this linear fit for extrapolating the configurational entropy (see Sec. E4).

### 3. Mixing entropy

The mixing entropy of the continuous polydisperse system formally diverges in the thermodynamic limit because each particle in the system then belongs to one of an infinite number of different components [24, 50]. One, however, can subtract from this quantity a physically relevant contribution, which we call the effective mixing entropy,  $s_{\text{mix}}^*$ . The main idea is that a continuous polydisperse system can be regarded as an effective  $M^*$ -components system (see Ref. [51] and references therein). We then assume that the effective  $M^*$ -components system shares physical properties, in particular a same (free-)energy basin, with the original continuous polydisperse system. Here, we provide a practical description of how to estimate  $s_{\text{mix}}^*$  from this scheme in our system; a full explanation is provided in Ref. [25].

To estimate  $M^*$ , we first decompose the distribution  $f(\sigma)$  into  $M$  species, as shown in the inset of Fig. S4(B). Here, we define  $M$  species by dividing  $f(\sigma)$  into equal intervals  $\Delta\sigma = (\sigma_{\max} - \sigma_{\min})/M$ , such that each species occupies the same fraction of the total volume,  $A\rho\pi\Delta\sigma/6 = C$ , where  $C$  is a constant. Note that  $M$  is an integer and that  $M \rightarrow \infty$  corresponds to the original continuous polydisperse system. We then perform a quench of the discretized system from the original configuration, and determine whether or not it remains in the same glassy basin as the original system by measuring the mean-squared displacement before and after the quench,

$$\Delta_M = \left[ \frac{1}{N} \sum_{i=1}^N |\mathbf{r}_{Mi}^{\text{IS}} - \mathbf{r}_{0i}|^2 \right], \quad (\text{E9})$$

where  $\{\mathbf{r}_{Mi}^{\text{IS}}\}$  is the inherent structure configuration of a  $M$ -discretized system. If  $M$  is large, the discretized system is almost identical with the original continuous polydisperse system and stays within the same basin, hence  $\Delta_M \simeq \Delta_{M \rightarrow \infty}$ . If  $M$  is small, however, discretization destroys the glassy basin of the original system. The system thus structurally rearranges into a different glassy basin, and  $\Delta_M \gg \Delta_{M \rightarrow \infty}$ . This determination is done here by considering inherent structures, which for hard-sphere systems correspond to an out-of-equilibrium compression up to jamming [52]. We determine  $M^*$  as the crossover between these two limit cases.

More precisely, we follow the following algorithm.

- 1) Obtain an equilibrium configuration of the original continuous polydisperse system,  $\{\mathbf{r}_{0i}\}$ , for the initial configuration.
- 2) Discretize the diameters  $\sigma$  of the original system into  $M$  species, keeping  $\phi$  fixed.
- 3) Quench the system to its inherent structure,  $\{\mathbf{r}_{Mi}^{\text{IS}}\}$ , using the algorithm described in Refs. [53, 54].
- 4) Repeat 1) - 3) for a range of  $M$  and over different initial configurations.
- 5) Determine the crossover  $M^*$  from  $\Delta_M$  as a function of  $M$ .

In Fig. S4(A), we show how  $\Delta_M$  evolves with  $M$ . At large  $M$ ,  $\Delta_M$  is flat and near  $\Delta_{M \rightarrow \infty}$ . The discretized system stays in its original basin. Upon decreasing  $M$ , however,  $\Delta_M$  deviates from  $\Delta_{M \rightarrow \infty}$ , indicating that the discretized system escapes its basin. We estimate  $M^*$  as an onset of this deviation by fitting the two linear lines from large and small  $M$  regions. We find a weak  $\phi$  dependence with  $M^* \approx 9 - 10$  for all  $\phi$  considered. We use linear fits to determine the more precise location of the crossover at  $M^* = 9.13$ , see Fig. S4(A).

We map  $M^*$  onto the effective mixing entropy  $s_{\text{mix}}^*$  by computing  $s_{\text{mix}}$  for different  $M$ ,

$$s_{\text{mix}}(M) = - \sum_{m=1}^M X_m \ln X_m, \quad (\text{E10})$$

$$X_m = \int_{\sigma_{\min} + (m-1)\Delta\sigma}^{\sigma_{\min} + m\Delta\sigma} d\sigma f(\sigma), \quad (\text{E11})$$

where  $X_m$  is the concentration of species  $m$ . Fig. S4(B) shows  $s_{\text{mix}}(M)$ . From the mapping indicated by the arrows, we obtain  $s_{\text{mix}}^* = 1.98$ . Notice that  $s_{\text{mix}}^*$  in Fig. S4(B) varies relatively weakly with  $M^*$  and so a more precise estimate of the value of  $M^*$  is not needed for our purposes.

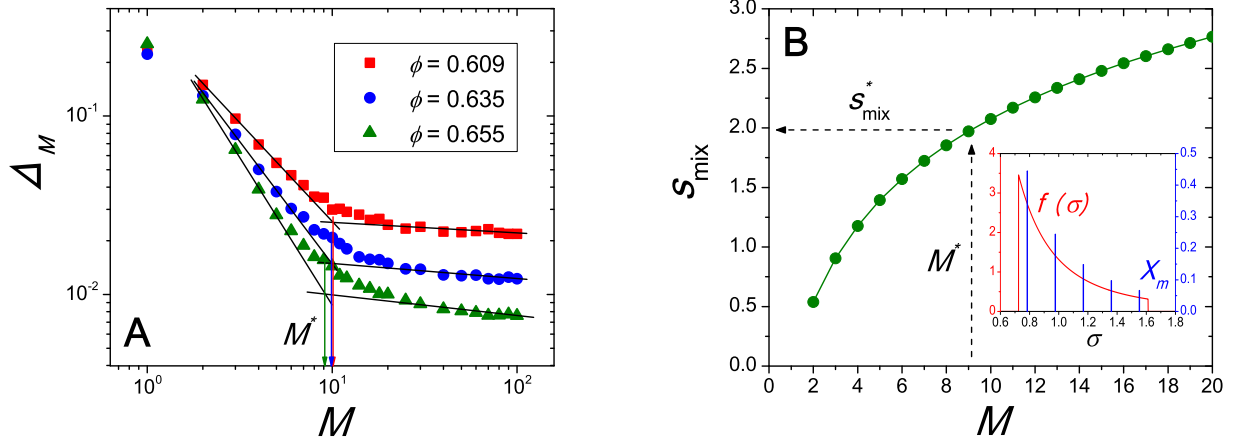


FIG. S4. (A) Mean-squared displacement  $\Delta_M$  between configurations before and after the quench. The crossover,  $M^*$ , is determined from the intersection of the two linear regimes at small and large  $M$ . (B) The mixing entropy  $s_{\text{mix}}$  as a function of  $M$  from Eqs. (E10) and (E11). The dashed arrows map  $M^*$  onto  $s_{\text{mix}}^*$ . The inset shows the distributions  $f(\sigma)$  and  $X_m$  for  $M = 5$ , as an example.

#### 4. Configurational entropy

Using the above results, we obtain the configurational entropy from  $s_{\text{tot}} - s_{\text{vib}}$  as a function of  $\phi$ . As can be seen in Fig. S5, no significant finite size effect is found between  $N = 1000$  and  $8000$ . We also find that the combination of  $s_{\text{tot}}$  from the (modified) BMCSL EOS in Fig. S3(A) and  $s_{\text{vib}}$  from the linear fit in Fig. S3(B) gives a reasonable extrapolation toward  $s_{\text{tot}} - s_{\text{vib}} = 0$ , and thus an estimate for the Kauzmann transition,  $\phi_K \approx 0.68$  ( $Z_K \approx 45$ ). In Fig. 3 from the main text we use the curve extrapolated by the modified BMCSL EOS.

#### Appendix F: Methods 2 and 3—The Franz-Parisi free energy

In this section we detail the Franz-Parisi construction [26]. We describe the order parameter of the glass transition, the umbrella sampling, as well as the parallel tempering and histogram reweighting techniques used to measure numerically the configurational entropy defined by Methods 2 and 3. Because the presentation closely follows earlier reports in which the Franz-Parisi free energy was determined numerically [11, 55, 56], we only provide a brief summary of the computational methodology used.

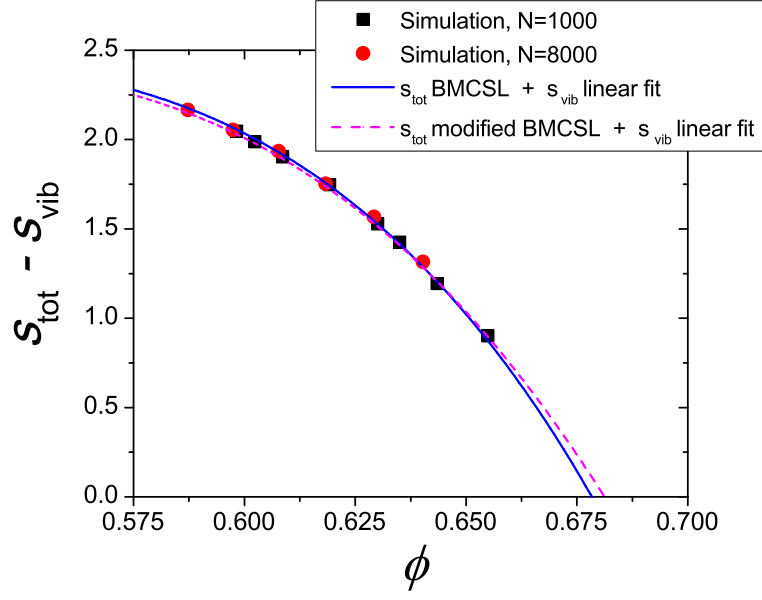


FIG. S5. Configurational entropy  $s_{\text{tot}} - s_{\text{vib}}$  as a function of volume fraction  $\phi$ , including the effective mixing entropy  $s_{\text{mix}} = s_{\text{mix}}^* = 1.98$ . Solid (dashed) curve is a combination of  $s_{\text{tot}}$  from the (modified) BMCSL EOS and  $s_{\text{vib}}$  from the linear fit.

### 1. Order parameter

The order parameter of the Franz-Parisi construction is the overlap  $Q$  (or similarity) between two disordered configurations, defined as

$$Q = \frac{1}{N} \sum_{i,j} \theta(a - |\mathbf{r}_{1,i} - \mathbf{r}_{2,j}|), \quad (\text{F1})$$

where  $\theta(x)$  is the Heaviside function, and  $\mathbf{r}_{1,i}$  and  $\mathbf{r}_{2,j}$  are the positions of particle  $i$  and  $j$  within configurations 1 and 2, respectively. By definition,  $Q$  is large when the density fields in configurations 1 and 2 are very close, and  $Q$  is close to zero when the density profiles are uncorrelated. The length,  $a = 0.23$ , accounts for small-amplitude thermal vibrations, so that density profiles that only differ by local vibrations caused by thermal fluctuations have  $Q$  close to unity.

### 2. Umbrella sampling

First, a reference configuration,  $\{\mathbf{r}_2\}$ , is chosen at random from the equilibrium states of the system at a fixed volume fraction. In order to access the probability distribution function of the overlap  $Q_{12}$  between configurations  $\{\mathbf{r}_1\}$  and  $\{\mathbf{r}_2\}$ , we conduct  $n$  distinct simulations. In each simulation,  $\ell = 1, \dots, n$ , system 1 evolves according to the Hamiltonian

$$H_\ell = H(\{\mathbf{r}_1\}) - \varepsilon Q_{12} + W_\ell(Q_{12}), \quad (\text{F2})$$

where  $H(\{\mathbf{r}\})$  is the original hard-sphere Hamiltonian and  $\varepsilon$  is the thermodynamic field conjugate to the overlap  $Q_{12}$ . The biasing potential  $W_\ell(Q)$  is taken to have the form

$$W_\ell(Q) = k_\ell (Q - Q_\ell)^2, \quad (\text{F3})$$

with parameters  $(k_\ell, Q_\ell)$  chosen to constrain the overlap  $Q_{12}$  to explore values away from its average equilibrium value. We then perform local displacements and swap MC moves, which are accepted using a Metropolis acceptance rate given by the Hamiltonian (F2).

Provided that the system is properly thermalized, our measurements yield the equilibrium probability distribution function of the overlap,

$$P_\ell(Q, \varepsilon, \phi) = \langle \delta(Q - Q_{12}) \rangle_\ell, \quad (\text{F4})$$

where  $\langle \dots \rangle_\ell$  denotes the thermal average with Hamiltonian  $H_\ell$  in Eq. (F2) at fixed reference configuration 2. Because we perform a quenched average to compute the overlap distribution, however, a second averaging, over the reference configuration 2, is needed to determine the Franz-Parisi free energy.

The idea behind the biasing potentials  $W_\ell(Q)$  in Eqs. (F2), and (F3) is that the fluctuations of the overlap in each simulation can be tailored to explore a narrow region centered around  $Q_\ell$ . Each simulation thus explores but a small range of overlap values, and it becomes unnecessary to wait for very rare overlap fluctuations to occur. Umbrella sampling enables the efficient measure of the atypical overlap fluctuations needed to reconstruct the Franz-Parisi free energy.

### 3. Parallel tempering

While the umbrella sampling technique described above considerably accelerates the measurement of the overlap fluctuations, we have observed that when  $\phi$  is high,  $N$  is large, and/or  $Q_\ell$  is large, the particle dynamics still slows down considerably. It then becomes difficult to perform an accurate sampling of the overlap fluctuations imposed by the Hamiltonian (F2), because the overlap fluctuations become slow.

To accurately sample the density regime explored in the present work, we have implemented parallel tempering MC moves [57]. We conduct the  $n$  simulations needed for the umbrella sampling at volume fraction  $\phi$  in parallel, and propose MC exchange moves between neighboring simulations, as characterized by nearby sets of parameters, say  $(k_\ell, Q_\ell)$  and  $(k_{\ell+1}, Q_{\ell+1})$ . An exchange between simulations  $\ell$  and  $\ell + 1$  is proposed with a low frequency (typically every  $10^3$  MC sweeps) and they are accepted with a Metropolis acceptance rate given by the Hamiltonians  $H_\ell$  and  $H_{\ell+1}$ , ensuring detailed balance.

Because each simulation now performs a random walk in the parameter space defined by  $\{(k_\ell, Q_\ell), \ell = 1, \dots, n\}$ , the sampling of overlap fluctuations is greatly enhanced. For the method to be efficient, however, we need to adjust the biasing potentials  $W_\ell(Q)$  such that the distributions  $P_\ell(Q)$  of neighboring simulations overlap significantly. We have thus used up to  $n = 20$  to gather data. Thermalization was carefully checked by running long simulations (up to  $10^{10}$  MC sweeps per simulation), in order to make sure that each state point was visited several times by all simulations via the replica exchange. This represents a significant numerical effort.

### 4. Histogram reweighting

Having obtained thermalized results from  $n$  biased simulations run in parallel, we process the simulation outcome using multi-histogram reweighting methods to reconstruct the unbiased probability  $P(Q)$  from the independently measured  $P_\ell(Q)$ ,

$$P(Q, \varepsilon, \phi) = \frac{\sum_{\ell=1}^n P_\ell(Q, \varepsilon, \phi)}{\sum_{\ell=1}^n e^{-\beta W_\ell} / Z_\ell}, \quad (\text{F5})$$

where  $Z_\ell$  are defined self-consistently as

$$Z_\ell = \int_0^1 dQ' \frac{\sum_{m=1}^n P_j(Q', \varepsilon, \phi)}{\sum_{m=1}^n e^{\beta(W_\ell - W_m)} / Z_m}. \quad (\text{F6})$$

Note that the value of  $\varepsilon$  used in the simulations plays no conceptual role because reweighting directly provides  $P(Q, \varepsilon', \phi)$  from  $P(Q, \varepsilon, \phi)$  for distinct field values  $\varepsilon$  and  $\varepsilon'$ :

$$P(Q, \varepsilon', \phi) = \frac{P(Q, \varepsilon, \phi) e^{-\beta Q(\varepsilon' - \varepsilon)}}{\int_0^1 dQ' P(Q', \varepsilon, \phi) e^{-\beta Q'(\varepsilon' - \varepsilon)}}. \quad (\text{F7})$$

Two values of the field  $\varepsilon$  are nonetheless particularly relevant to determining the configurational entropy, as described in the following two subsections.

## 5. Method 2

First, the Franz-Parisi free energy  $V(Q)$  is obtained as

$$V(Q) = -\lim_{\varepsilon \rightarrow 0} \left[ \frac{T}{N} \ln P(Q, \varepsilon, \phi) \right], \quad (\text{F8})$$

where  $[\dots]$  denotes averaging over the quenched reference configuration 2. We have used 60 independent reference configurations for each value of the volume fraction. Because  $V(Q)$  is only defined up to an additive constant, we arbitrarily adjust it, such that  $V(Q_{\text{low}}) = 0$ , where  $Q_{\text{low}} \approx 0.05$  is defined as the location of the global minimum in  $V(Q)$ . This additive constant is irrelevant because only the free energy difference,

$$s_{\text{conf}} = V(Q_{\text{high}}) - V(Q_{\text{low}}) \quad (\text{Method 2}) \quad (\text{F9})$$

is needed to determine the configurational entropy. In this work, we set  $Q_{\text{high}} = 0.8$  as motivated in the following subsection.

## 6. Method 3

Second, we obtain the critical field value  $\varepsilon = \varepsilon^*$  needed to induce a phase coexistence between low and high overlap states [26, 56]. In practice, we use the strength of the reweighting method to finely explore a range of  $\varepsilon$  values, and define  $\varepsilon^*$  as the field strength for which the overlap shows a value intermediate between  $Q_{\text{low}}$  and  $Q_{\text{high}}$ , where the distribution  $P(Q, \varepsilon^*, \phi)$  shows two peaks of equal amplitude, and where the variance of the overlap fluctuations (the susceptibility) is maximal. The position of the second peak of  $P(Q, \varepsilon^*, \phi)$  at coexistence sets  $Q_{\text{high}}$ . Because its volume fraction dependence is very weak, we fix it close to its average value and use  $Q_{\text{high}} = 0.8$  for all  $\phi$ . Note that a secondary minimum cannot exist in the large system size limit of finite-dimensional simulations [11], which is why we resort to the above definition of  $Q_{\text{high}}$ .

The field  $\varepsilon^*$  has a simple graphical interpretation. It represents the amplitude of the field needed to ‘tilt’ the potential  $V(Q)$  towards coexistence [see Fig. 2(B) of the main text]. Because the relation  $V(Q_{\text{high}}) \approx \varepsilon^*(Q_{\text{high}} - Q_{\text{low}})$ , holds to a good approximation,  $\varepsilon^*$  provides an estimate of the free energy difference  $V(Q_{\text{high}}) - V(Q_{\text{low}})$ . We can thus define the configurational entropy as

$$s_{\text{conf}} = \varepsilon^*(Q_{\text{high}} - Q_{\text{low}}) \quad (\text{Method 3}). \quad (\text{F10})$$

## Appendix G: Method 4—The point-to-set correlation

### 1. Definition of cavity core overlap and point-to-set correlations

The similarity between two configurations,  $\mathbf{X} = \{\mathbf{x}_i\}$  and  $\mathbf{Y} = \{\mathbf{y}_i\}$ , is characterized by using an overlap field,  $q_{\mathbf{X}, \mathbf{Y}}(\mathbf{r})$ , computed as in Ref. [20]. For each particle  $\mathbf{x}_i$ , we find the nearest particle  $\mathbf{y}_{i_{\text{nn}}}$  and assign an overlap value  $q_{\mathbf{X}, \mathbf{Y}}(\mathbf{x}_i) \equiv w(|\mathbf{x}_i - \mathbf{y}_{i_{\text{nn}}}|)$ , where

$$w(z) \equiv \exp \left[ - \left( \frac{z}{b} \right)^2 \right], \quad (\text{G1})$$

with  $b = 0.2$ . This function defines overlap values  $q_{\mathbf{X}, \mathbf{Y}}(\mathbf{x}_i)$  at scattered points  $\{\mathbf{x}_i\}$ . We then define a continuous function passing through these points,  $q_{\mathbf{X}, \mathbf{Y}}(\mathbf{r})$ . Specifically, we first perform a Delaunay tessellation of space and, to a point  $\mathbf{r}$  within a simplex spanned by four points  $\{\mathbf{x}_i\}_{i=i_1, i_2, i_3, i_4}$ , associate a linearly interpolated value

$$q_{\mathbf{X}, \mathbf{Y}}(\mathbf{r}) = \sum_{i=i_1, i_2, i_3, i_4} c_i q_{\mathbf{X}}(\mathbf{x}_i), \quad (\text{G2})$$

where  $\{c_i\}_{i=i_1, i_2, i_3, i_4}$  satisfies  $\mathbf{r} = \sum_{i=i_1, i_2, i_3, i_4} c_i \mathbf{x}_i$  with the constraint  $\sum_{i=i_1, i_2, i_3, i_4} c_i = 1$ . We can similarly obtain  $q_{\mathbf{Y}, \mathbf{X}}(\mathbf{r})$ , allowing us to define the overlap field

$$q_{\mathbf{X}, \mathbf{Y}}(\mathbf{r}) \equiv \frac{1}{2} \{q_{\mathbf{X}, \mathbf{Y}}(\mathbf{r}) + q_{\mathbf{Y}, \mathbf{X}}(\mathbf{r})\}. \quad (\text{G3})$$

In order to capture the similarity between two configurations near the center of the cavity, we also define the cavity core overlap

$$q_c \equiv \frac{3}{4\pi r_c^3} \int_{|\mathbf{r}'| < r_c} d\mathbf{r}' q_{\mathbf{X}, \mathbf{Y}}(\mathbf{r}'), \quad (\text{G4})$$

where  $r_c = 0.5$  and  $\mathbf{r}' = \mathbf{0}$  is the cavity center. This integral is numerically evaluated by Monte Carlo integration with  $10^4$  points.

For each volume fraction  $\phi$  and cavity radius  $R$ , the so-called point-to-set correlation function,  $Q_{\text{PTS}}(R; \phi)$ , is evaluated by disorder-averaging over 40 cavity centers and, within each cavity, thermal-averaging over  $s_{\text{prod}}$  pairs of equilibrated configurations (see Tables IV-X) [20]. We extract the point-to-set correlation length through the compressed exponential fit,

$$Q_{\text{PTS}}(R; \phi) = A \exp[-\{R/\xi_{\text{PTS}}(\phi)\}^\gamma] + Q_{\text{PTS}}^{\text{bulk}}(\phi), \quad (\text{G5})$$

with the compressed exponent  $\gamma = 4$ . Here the bulk value,  $Q_{\text{PTS}}^{\text{bulk}}$ , corresponds to the value at  $R = \infty$  and is evaluated by taking  $10^5$  pairs of independent configurations in bulk samples.

At the point-to-set length scale, we expect the probability distribution function of core overlaps to display broad fluctuations. Fig. S6 bears out this expectation. In particular, as anticipated in Ref. [20], we observe that the full disorder-averaged distribution becomes bimodal at high volume fractions, showing that we have indeed entered the deeply glassy regime that remains currently inaccessible for a thermal binary Lennard-Jones liquid.

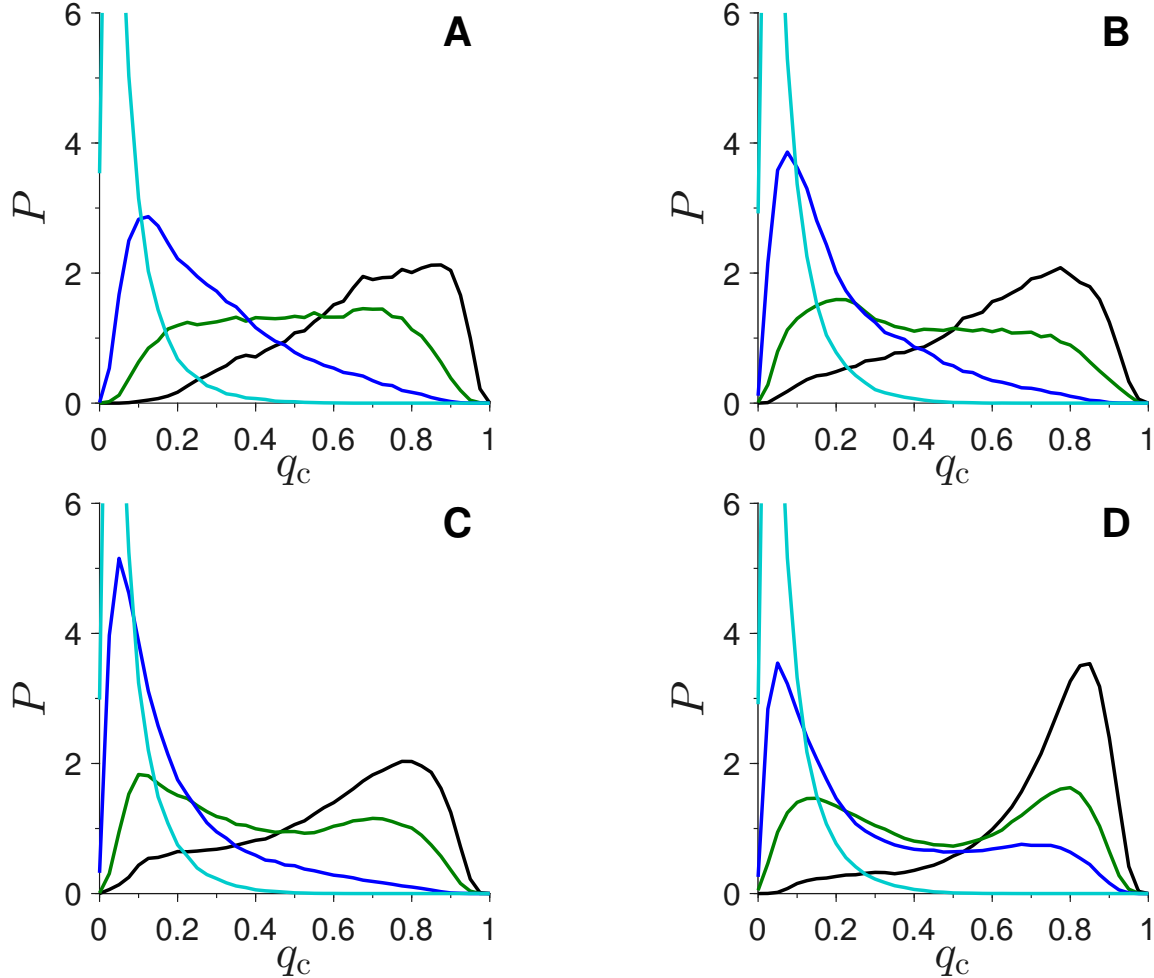


FIG. S6. Disorder-averaged probability distribution function of core overlap  $P(q_c)$ , at volume fractions  $\phi = 0.568$  (A),  $0.597$  (B),  $0.618$  (C), and  $0.640$  (D) with varying radius  $R = R_1$  (black),  $R_2$  (green),  $R_3$  (blue), and  $R_4$  (cyan). We chose  $(R_1, R_2, R_3, R_4)|_{\phi=0.568} = (1.2, 1.5, 1.8, 2.7)$ ;  $(R_1, R_2, R_3, R_4)|_{\phi=0.597} = (1.65, 1.95, 2.25, 3.0)$ ;  $(R_1, R_2, R_3, R_4)|_{\phi=0.618} = (1.95, 2.25, 2.7, 3.6)$ ; and  $(R_1, R_2, R_3, R_4)|_{\phi=0.640} = (2.4, 3.0, 3.3, 4.5)$ .



## 2. Parallel-tempering sampling with fuzzy-ensemble replicas

Denote positions of mobile particles inside the cavity  $\mathbf{x}_i^M$  with  $i = 1, \dots, N_{\text{cav}}$  and their associated diameters  $\sigma_i^M$ . Similarly, denote positions and diameters of pinned particles outside the cavity  $\mathbf{x}_j^P$  and  $\sigma_j^P$ , respectively. For cavity point-to-set measurements, one wishes to sample cavity configurations  $\{\mathbf{x}_i^M\}$  with probability

$$P(\{\mathbf{x}_i^M\} | \{\sigma_i^M\}, \{\mathbf{x}_j^P\}, \{\sigma_j^P\}) \quad (\text{G6})$$

$$\simeq \exp[-\beta U(\{\mathbf{x}_i^M\}, \{\sigma_i^M\}, \{\mathbf{x}_j^P\}, \{\sigma_j^P\})],$$

where  $\simeq$  denotes equality up to a normalization constant and a hard wall constraint  $|\mathbf{x}_i^M| < R$  at the edge of the cavity of size  $R$  is imposed. Note that for the standard hard-sphere potential,  $U$ , this expression does not depend on the inverse temperature  $\beta \in (0, \infty)$ . Properly sampling from this weight can be computationally demanding, because the equilibration time grows rapidly both as the cavity size decreases and as the volume fraction increases (see Fig. S7 and Sec. G 4). In these systems, phase space can actually become fully disconnected, which makes proper sampling impossible through (semi-)local Monte Carlo moves.

We here overcome this difficulty through a parallel-tempering scheme [38, 58]. The simplest scheme, analogous to that employed in Ref. [20] for a thermal binary Lennard-Jones liquid, would be to couple the above ensemble with replicas governed by a probability

$$P_\lambda(\{\mathbf{x}_i^M\} | \{\sigma_i^M\}, \{\mathbf{x}_j^P\}, \{\sigma_j^P\}) \quad (\text{G7})$$

$$\simeq \exp[-\beta U(\{\mathbf{x}_i^M\}, \{\lambda \sigma_i^M\}, \{\mathbf{x}_j^P\}, \{\sigma_j^P\})],$$

where the diameter of the particles inside the cavity is uniformly shrunk by  $\lambda$ . This approach, however, is highly inefficient for hard interactions. Replica-swap attempts are indeed very likely to be rejected, because even small  $\lambda$  increments can result in hard overlaps. More precisely, in order to have appreciable replica-swapping acceptance rates, the spacing between consecutive  $\lambda$ 's must scale as  $O(1/N_{\text{cav}})$ , and thus  $O(N_{\text{cav}})$  replicas are needed, as opposed to  $O(\sqrt{N_{\text{cav}}})$  for thermal systems [38].

We therefore introduce a *fuzzy* ensemble governed by probability weights controlled by two parameters,  $\alpha$  and  $\hat{\lambda}$ ,

$$P_{\alpha, \hat{\lambda}}(\{\mathbf{x}_i^M\}, \lambda | \{\sigma_i^M\}, \{\mathbf{x}_j^P\}, \{\sigma_j^P\}) \quad (\text{G8})$$

$$\simeq \exp\left\{-\frac{\alpha^2 N_{\text{cav}}}{2} (\lambda - \hat{\lambda})^2\right\}$$

$$\times \exp[-\beta U(\{\mathbf{x}_i^M\}, \{\lambda \sigma_i^M\}, \{\mathbf{x}_j^P\}, \{\sigma_j^P\})],$$

hence the shrinking factor  $\lambda$  is allowed to fluctuate. With this modification, a  $O(\sqrt{N_{\text{cav}}})$  scaling for the number of replicas is recovered. One caveat is that one of the replicas must obey the original sharp distribution defined in Eq. (G6)—which formally corresponds to the limit  $\alpha \rightarrow \infty$  with  $\hat{\lambda} = 1$ —and the aforementioned challenge imposed by hard interactions still persists for that case. This problem is here surmounted by using a large number of replica-swap attempts near the original replica [order  $O(1000 N_{\text{cav}})$  more frequent than others], as detailed in Sec. G 3 b.

## 3. Details of Monte Carlo moves

This section details the implementation of the Monte Carlo scheme for the cavity simulations.

### a. Basic moves within a replica

Within each replica, we perform Monte Carlo moves consisting of local displacements, particle identity swaps, and, for fuzzy ensembles, diameter fluctuations.

- Local displacements consist of: (i) randomly choosing a particle  $i$  from the  $N_{\text{cav}}$  mobile particles within the cavity; (ii) displacing particle  $i$  by  $\Delta \mathbf{x} = l \hat{\mathbf{n}}$  with uniform  $l \in [0, 0.088]$  and uniform  $\hat{\mathbf{n}}$  on the unit sphere  $S^2$ ; and (iii) accepting displacement only if no hard overlaps are created. Note that an additional hard spherical wall at the edge of the cavity guarantees that mobile particle cannot leave the cavity—a rare instance, even at the lowest volume fraction studied.

- Particle identity swaps consist of: (i) choosing two distinct particles  $i$  and  $j$  within a cavity; and (ii) swapping their diameters if it results in no hard overlaps. (For the highest volume fraction  $\phi = 0.640$  with the cavity size  $R \geq 2.7$ , in order to accelerate runs, we attempt these moves only for pairs with the diameter difference  $\lambda|\sigma_i - \sigma_j| < 0.075$ .)
- Diameter fluctuations consist of: (i) uniformly drawing a shift  $\Delta\lambda \in \left[-\frac{0.2}{\alpha\sqrt{N_{\text{cav}}}}, +\frac{0.2}{\alpha\sqrt{N_{\text{cav}}}}\right]$ , and (ii) accepting the shift with probability  $p = \min\{p_\Lambda, 1\}$ , where

$$p_\Lambda = \exp \left[ -\frac{\alpha^2 N_{\text{cav}}}{2} \left\{ \left( \lambda + \Delta\lambda - \hat{\lambda} \right)^2 - \left( \lambda - \hat{\lambda} \right)^2 \right\} \right], \quad (\text{G9})$$

if no hard overlaps are created.

One MC sweep consists of  $N_{\text{cav}}$  MC trial moves of the above kinds for each replica, following the composition detailed in the next subsection.

### b. Parallel tempering

For large cavities ( $R > \xi_{\text{PTS}} + 1$ ), each MC sweep then consists of  $N_{\text{cav}}$  MC trial moves with 70% (80%) local displacements and 30% (20%) particles identity swaps (parenthesis are for  $\phi = 0.640$ ), but diameter fluctuations and replica exchange are not needed (see Fig. S7). For small cavities, however, proper sampling requires replica exchange.

We label replicas  $a = 1, \dots, n$ , where  $a = 1$  corresponds to the original ensemble with  $\hat{\lambda}_1 = 1$  and  $\alpha_1 = \infty$ . For the other replicas, we choose  $\alpha_{a \geq 2} = 20$ , and  $\{\hat{\lambda}_a\}$  are tuned to enable appreciable replica-swap rate (see below). For  $a \geq 3$ , each MC sweep consists of  $N_{\text{cav}}$  MC trial moves with 60% (70%) local displacements, 30% (20%) particles identity swaps, and 10% diameter fluctuations (parenthesis are for  $\phi = 0.640$  with  $R \geq 2.7$ ). For  $a = 1$  and  $a = 2$ , a different scheme is employed. The two cavities are run as a pair with a large number of replica-swap attempts: each MC sweep consists of  $4N_{\text{cav}}$  attempts, 50% replica-swap between  $\{\mathbf{x}_i^{\text{M}}\}_1$  and  $\{\mathbf{x}_i^{\text{M}}\}_2$  –accepted only when swapping  $\lambda_1 = 1$  and  $\lambda_2$  does not result in hard overlaps–, 15% (17.5%) local displacements respectively for  $a = 1$  and  $a = 2$ , 7.5% (5%) particles identity swaps respectively for  $a = 1$  and  $a = 2$ , and 5% diameter fluctuations for  $a = 2$ .

For each  $a \geq 2$ , a replica-identity swap between  $(\{\mathbf{x}_i^{\text{M}}\}_a, \lambda_a)$  and  $(\{\mathbf{x}_i^{\text{M}}\}_{a+1}, \lambda_{a+1})$  is attempted every 1000 MC sweeps on average, with acceptance probability  $p = \min\{p_{\text{RS}}, 1\}$ , where

$$p_{\text{RS}} = \exp \left\{ -\alpha^2 N_{\text{cav}} \left( \hat{\lambda}_{a+1} - \hat{\lambda}_a \right) (\lambda_{a+1} - \lambda_a) \right\}. \quad (\text{G10})$$

The replica parameters,  $\{\hat{\lambda}_a\}_{a \geq 2}$ , are tuned to ensure sufficient replica-swap rates. In order to achieve this sampling, we first define the average of the fluctuating  $\lambda$

$$\langle \lambda \rangle_a = \int d\lambda d\mathbf{x}_i^{\text{M}} \lambda \mathbf{P}_{\alpha_a, \hat{\lambda}_a}(\{\mathbf{x}_i^{\text{M}}\}_a, \lambda \mid \{\sigma_i^{\text{M}}\}, \{\mathbf{x}_j^{\text{P}}\}, \{\sigma_j^{\text{P}}\}), \quad (\text{G11})$$

For the replica  $a = 2$ , we ensure that  $\langle \lambda \rangle_2 = 1 + O\left(\sqrt{\langle \lambda^2 \rangle_2 - \langle \lambda \rangle_2^2}\right)$ , which typically requires  $\hat{\lambda}_2 > 1$ , because of the system's relatively high pressure. Replicas are then added one by one, with  $\hat{\lambda}_2 > \hat{\lambda}_3 > \dots > \hat{\lambda}_n$ , each time targeting a replica-swap acceptance rate of  $\sim 20\%$ . This process is stopped upon  $\hat{\lambda}_n$ , such that  $\langle \lambda \rangle_n < \lambda_{\text{dec}}$  (see Tables IV-X). Although this linear approach does not attain a globally uniform replica-swap acceptance rate (see [59, 60] for more systematic approaches), the resulting scheme was nonetheless sufficient to ensure equilibration and convergence, as defined in the next subsection.

The concerted use of the replica exchange, the fuzzy ensembles, and the specialized sampling scheme around the original ensemble makes possible to access a sufficient number of independent cavity configurations from the desired ensemble defined in Eq. (G6).

### c. Convergence criterion

The quality of equilibration within each cavity is evaluated by monitoring the convergence of two schemes [20, 61]: (i) starting sampling from the original configuration, and (ii) starting sampling from a randomized configuration

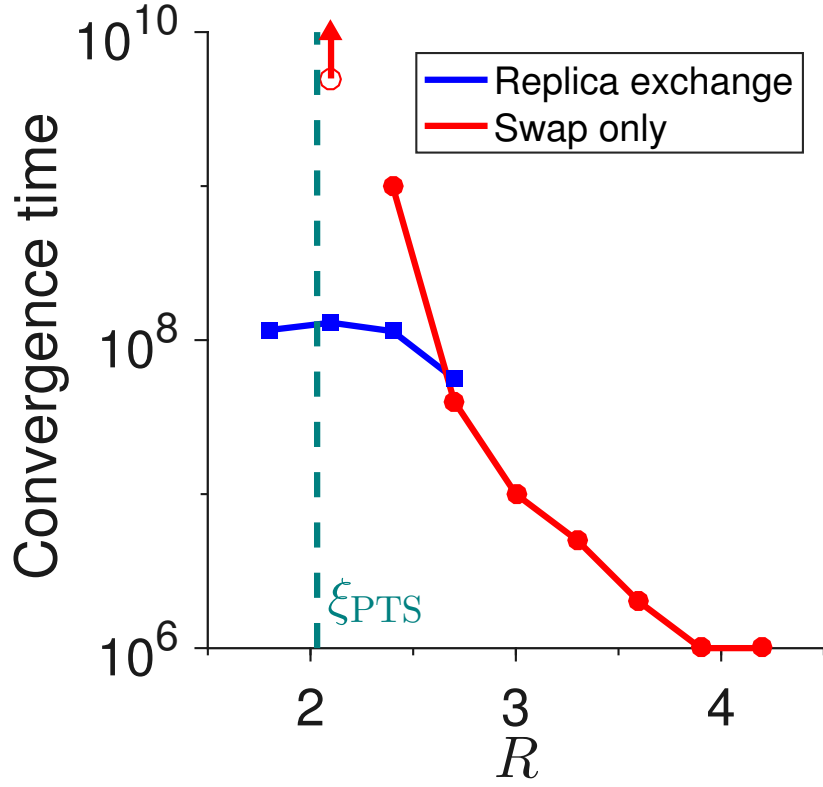


FIG. S7. Characteristic relaxation time (in unit of MC sweeps) multiplied by the number of replicas with (blue-square) and without (red-circle) replica exchange, as a function of cavity size  $R$ . Note that the result without replica exchange at  $R = 2.1$  is an underestimate, because particle swaps alone did not allow the results to converge even after  $0.5 \cdot 10^{10}$  MC sweeps. We also checked that convergence is not achieved for  $R = 1.5$  (not shown).

prepared by first running  $10^6$  MC sweeps with shrunk cavity particles at  $(\hat{\lambda}, \alpha) = (0.5, 20)$  and then slowly regrowing them back to  $\lambda = 1$ . We then record configurations every  $t_{\text{rec}}$  MC sweeps (see Tables IV-X), and monitor the core overlap between new configurations and the original configuration,  $q_c^{\text{on}}(t)$ , as a function of the number of MC sweeps,  $t$ . The first  $s_{\text{eq}}$  configurations are discarded, and the average overlap for the following  $s_{\text{prod}}$ ,

$$\langle q_c^{\text{on}} \rangle \equiv \frac{1}{s_{\text{prod}}} \sum_{s=s_{\text{eq}}+1}^{s_{\text{eq}}+s_{\text{prod}}} q_c^{\text{on}}(t_{\text{rec}}s) \quad (\text{G12})$$

is computed. Convergence is deemed achieved when the results of both approaches lie within  $\pm 0.1$  of each other for each cavity. This criterion also allows us to estimate the convergence time for a given value of  $R$  and  $\phi$ . Replica parameters as well as  $s_{\text{eq}}$  and  $s_{\text{prod}}$  (see Tables IV-X) are chosen, such that at least 95% out of 40 cavities pass this convergence test, except the most challenging data point,  $(\phi, R) = (0.640, 3.3)$ , where a 92.5% rate was tolerated. Because the difference between the two approaches is not systematic, however, averaging over 40 cavities results in a rather close agreement between the two schemes, *i.e.*, a convergence to within  $\pm 0.01$ .

#### 4. Swap is not enough

It has been suggested that particle identity swaps by themselves suffice to thermalize cavities, especially small ones with  $R < \xi_{\text{PTS}}$  [29, 61]. We here test this hypothesis for our polydisperse system, for which particle swaps are extremely effective at sampling bulk configurations. Fig. S7 contrasts the computational time needed to equilibrate cavity configurations at  $\phi = 0.608$ , with and without parallel tempering. Without replica exchange, the equilibration time rapidly grows as the cavity size decreases, which is in line with the physical suggestion that confinement enhances the breaking of ergodicity [20]. We observe that when  $R$  approaches the point-to-set length from above, however, the convergence time becomes too long to be measured. Tests performed with even smaller  $R$  values confirm that trend.

Convergence is never achieved for cavities of the order of the point-to-set length or smaller. By contrast, using replica exchange with swap dynamics allows to maintain a reasonable convergence time even for small cavities.

This result seems to contrast with the findings of Ref. [61] that the dynamics actually speeds up inside small cavities. We have no explanation for this discrepancy as we never observed the reported speed-up phenomenon in any of our simulations, but we note that the timescale studied in Ref. [61] is different from the one shown in Fig. S7. Specifically, whereas we report a convergence time for the overlap inside the cavity, Cavagna *et al.* study the time decay of the overlap-*fluctuation* auto-correlation function in equilibrium. We have not tried to systematically measure this latter auto-correlation because it is already very clear that, according to the former convergence criterion, we cannot properly thermalize small cavities for simulations with swaps alone. Therefore we would not access an equilibrium correlation timescale but a nonequilibrium one. In addition, no quantitative analysis of the convergence time is provided in Refs. [29, 61], which makes a direct comparison with our results impossible.

We conclude that one must generally employ a parallel-tempering scheme, as developed in Ref. [20], in order to properly sample small cavities and measure point-to-set correlations.

$R$	1.2	1.5	1.8	2.1	2.4	2.7	3.0	3.3	3.6
$n_{\text{ave}}$	10	8	7	7	6	1	1	1	1
$\lambda_{\text{dec}}$	0.850	0.920	0.960	0.970	0.980	NA	NA	NA	NA
$t_{\text{rec}}$	4000	4000	4000	4000	4000	1000	400	200	200
$s_{\text{eq}}$	500	500	500	500	500	500	500	500	500
$s_{\text{prod}}$	2000	2000	2000	2000	2000	2000	2000	2000	2000

TABLE IV. Cavity PTS measurement parameters  $\phi = 0.568$ . Runs without parallel tempering have  $n_{\text{ave}} = 1$ .

$R$	1.5	1.8	2.1	2.4	2.7	3.0	3.3	3.6	3.9
$n_{\text{ave}}$	10	10	10	8	7	1	1	1	1
$\lambda_{\text{dec}}$	0.900	0.930	0.950	0.970	0.980	NA	NA	NA	NA
$t_{\text{rec}}$	$2 \cdot 10^4$	$2 \cdot 10^4$	$2 \cdot 10^4$	$10^4$	4000	2000	1000	400	400
$s_{\text{eq}}$	500	500	500	500	500	500	500	500	500
$s_{\text{prod}}$	2000	2000	2000	2000	2000	2000	2000	2000	2000

TABLE V. Cavity PTS measurement parameters  $\phi = 0.587$ . Runs without parallel tempering have  $n_{\text{ave}} = 1$ .

$R$	1.65	1.95	2.25	2.55	2.7	3.0	3.3	3.6	3.9
$n_{\text{ave}}$	12	13	12	11	9	1	1	1	1
$\lambda_{\text{dec}}$	0.900	0.920	0.940	0.960	0.970	NA	NA	NA	NA
$t_{\text{rec}}$	$2 \cdot 10^4$	$2 \cdot 10^4$	$2 \cdot 10^4$	$10^4$	$10^4$	$10^4$	4000	2000	1000
$s_{\text{eq}}$	500	500	500	500	500	500	500	500	500
$s_{\text{prod}}$	2000	2000	2000	2000	2000	2000	2000	2000	2000

TABLE VI. Cavity PTS measurement parameters  $\phi = 0.597$ . Runs without parallel tempering have  $n_{\text{ave}} = 1$ .

## Appendix H: Method 1 for soft spheres

We compute the configurational entropy of a continuous polydisperse soft sphere system by the standard definition,

$$s_{\text{conf}} = s_{\text{tot}} - s_{\text{vib}}, \quad (\text{H1})$$

where  $s_{\text{tot}}$  and  $s_{\text{vib}}$  are the total and vibrational entropies [10, 49]. The mixing entropy is also treated by the strategy in Ref. [25]; we compute  $s_{\text{tot}}$  and  $s_{\text{vib}}$  as if the system were monodisperse, then we add the effective mixing entropy  $s_{\text{mix}}^*$  which is independently determined by simulations [25].

$R$	1.8	2.1	2.4	2.7	3.0	3.3	3.6	3.9	4.2
$n_{\text{ave}}$	13	14	12	12	1	1	1	1	1
$\lambda_{\text{dec}}$	0.910	0.920	0.950	0.960	NA	NA	NA	NA	NA
$t_{\text{rec}}$	$2 \cdot 10^4$	$2 \cdot 10^4$	$2 \cdot 10^4$	$10^4$	$2 \cdot 10^4$	$10^4$	4000	2000	2000
$s_{\text{eq}}$	500	500	500	500	500	500	500	500	500
$s_{\text{prod}}$	2000	2000	2000	2000	2000	2000	2000	2000	2000

TABLE VII. Cavity PTS measurement parameters  $\phi = 0.608$ . Runs without parallel tempering have  $n_{\text{ave}} = 1$ .

$R$	1.95	2.1	2.25	2.7	3.3	3.6	3.9	4.2
$n_{\text{ave}}$	15	16	18	17	1	1	1	1
$\lambda_{\text{dec}}$	0.900	0.910	0.910	0.940	NA	NA	NA	NA
$t_{\text{rec}}$	$10^4$	$10^4$	$10^4$	$10^4$	$5 \cdot 10^4$	$3 \cdot 10^4$	$10^4$	5000
$s_{\text{eq}}$	2000	2000	2000	1000	500	500	500	500
$s_{\text{prod}}$	8000	8000	8000	4000	2000	2000	2000	2000

TABLE VIII. Cavity PTS measurement parameters  $\phi = 0.618$ . Runs without parallel tempering have  $n_{\text{ave}} = 1$ .

### 1. Model

We use a continuous size polydispersity, where the particle diameter  $\sigma$  of each particle is randomly drawn from the following particle size distribution:  $f(\sigma) = A\sigma^{-3}$ , for  $\sigma \in [\sigma_{\min}, \sigma_{\max}]$ , where  $A$  is a normalization constant. We define the size polydispersity as  $\Delta = \sqrt{\sigma^2 - \bar{\sigma}^2}/\bar{\sigma}$ , where  $\bar{\sigma} = \int d\sigma f(\sigma)(\sigma)$ . We use  $\Delta = 0.23$ , choosing  $\sigma_{\min}/\sigma_{\max} = 0.4492$ . We use  $\bar{\sigma}$  as the unit length. We simulate systems composed of  $N$  particles in a cubic cell of volume  $V$  with periodic boundary conditions in three dimensions,  $d = 3$ . We use the following pairwise potential for the soft sphere model,

$$v_{ij}(r) = v_0 \left( \frac{\sigma_{ij}}{r} \right)^n + c_0 + c_1 \left( \frac{r}{\sigma_{ij}} \right)^2 + c_2 \left( \frac{r}{\sigma_{ij}} \right)^4, \quad (\text{H2})$$

$$\sigma_{ij} = \frac{(\sigma_i + \sigma_j)}{2} (1 - \epsilon |\sigma_i - \sigma_j|), \quad (\text{H3})$$

where  $v_0$  is the unit of energy, and  $\epsilon$  quantifies the degree of non-additivity of the particle diameters. Non-additivity is introduced for convenience, as it prevents more efficiently crystallization and thus it enhances glass-forming ability of the numerical models. The constants,  $c_0$ ,  $c_1$  and  $c_2$ , are chosen so that the first and second derivatives of  $v_{ij}(r)$  become zero at the cut-off  $r_{\text{cut}} = 1.25\sigma_{ij}$ . We employ the non-additive soft sphere model using the parameters  $n = 12$  and  $\epsilon = 0.2$ . We set the number density  $\rho = N/V = 1.0$  with  $N = 1500$ .

Equilibrium configurations are produced by swap MC simulations where high temperature configurations are instantaneously quenched to the target temperature. Equilibration is ensured by the fact that particles lose memory of their initial positions and dynamical observables do not present aging. The absence of crystalline nuclei is inspected as well. We additionally perform standard MC simulations to obtain relaxation times down to the mode coupling crossover temperature  $T_c$ . Following the procedure described in the Materials and methods section of the main text, we carry out three kind of fits on these dynamical data. First we employ a power-law fit to extrapolate  $T_c = 0.104$ . Then we perform a VFT fit with  $\delta = 1$  and a parabolic fit to get the glass-ceiling temperatures  $T_g = 0.0720 - 0.0817$ .

$R$	2.4	2.55	2.7	3.0	3.3	3.9	4.2
$n_{\text{ave}}$	17	19	18	17	16	1	1
$\lambda_{\text{dec}}$	0.930	0.930	0.940	0.955	0.965	NA	NA
$t_{\text{rec}}$	$10^4$	$10^4$	$10^4$	$10^4$	$10^4$	$10^5$	$4 \cdot 10^4$
$s_{\text{eq}}$	1500	3000	4000	6000	8000	500	500
$s_{\text{prod}}$	6000	12000	16000	9000	12000	2000	2000

TABLE IX. Cavity PTS measurement parameters  $\phi = 0.629$ . Runs without parallel tempering have  $n_{\text{ave}} = 1$ .

$R$	2.4	2.7	3.0	3.15	3.3	4.2	4.5
$n_{\text{ave}}$	21	21	22	23	22	1	1
$\lambda_{\text{dec}}$	0.910	0.930	0.940	0.940	0.950	NA	NA
$t_{\text{rec}}$	$10^4$	$10^4$	$10^4$	$10^4$	$10^4$	$2 \times 10^5$	$10^5$
$s_{\text{eq}}$	2000	6000	8000	12000	12000	1000	1000
$s_{\text{prod}}$	8000	9000	12000	18000	18000	4000	4000

TABLE X. Cavity PTS measurement parameters  $\phi = 0.640$ . Runs without parallel tempering have  $n_{\text{ave}} = 1$ .

## 2. Total entropy

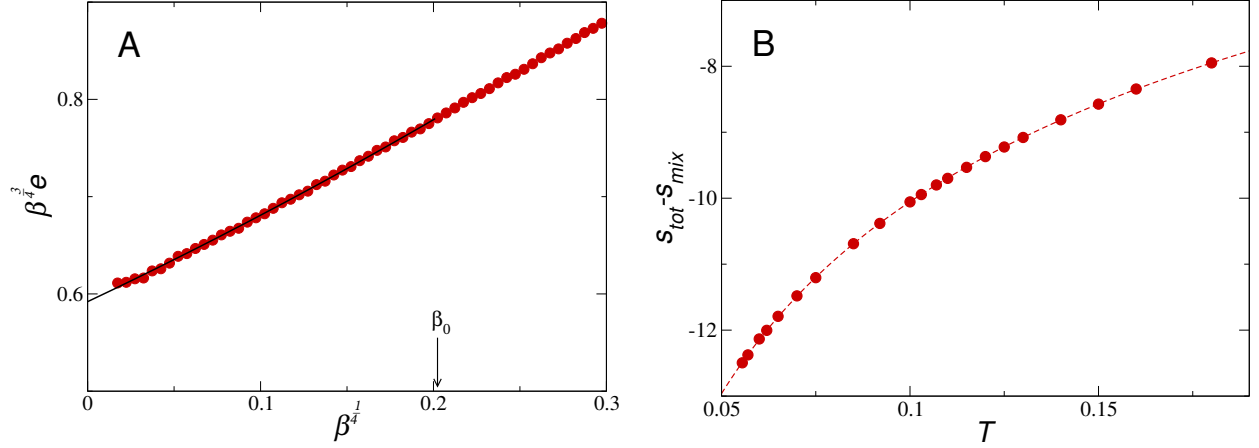


FIG. S8. (A)  $\beta^{3/4}e(\beta)$  vs.  $\beta^{1/4}$  plot to obtain the fitting parameters,  $A$ ,  $B$ , and  $C$ , in Eq. (H9). The solid line is the obtained fitting curve. The vertical arrow indicates the position of  $\beta_0$ . (B)  $s_{\text{tot}} - s_{\text{mix}}$  as a function of  $T$ . The dashed curve is the Rosenfeld-Tarazona expression,  $s_{\text{tot}} - s_{\text{mix}} + 3 \ln \Lambda \propto T^{-2/5}$ , which is well confirmed by our data.

We perform the thermodynamic integration over the inverse temperature  $\beta'$  from the ideal gas limit  $\beta' \rightarrow 0$  to the target temperature  $\beta$ ,

$$s_{\text{tot}}(\beta) = \frac{5}{2} - \ln \rho - 3 \ln \Lambda + \beta e(\beta) - \int_0^\beta d\beta' e(\beta') + s_{\text{mix}}, \quad (\text{H4})$$

where  $\Lambda = \sqrt{2\pi\beta\hbar^2/m}$  and  $e(\beta)$  is the potential energy of the system. We set  $m = 1$  and  $\hbar = 1$  for this system.

Special care is needed to compute the integral because the potential energy  $e(\beta)$  diverges in the high temperature limit [62, 63]. To accurately calculate the integral, we decompose the integration range into very high temperature regime,  $\beta' \in [0, \beta_0]$ , and the intermediate regime,  $\beta' \in (\beta_0, \beta]$ , by introducing the boundary  $\beta_0$  between two regimes. Therefore, the integral in Eq. (H4) can be decomposed as

$$I = \int_0^\beta d\beta' e(\beta') = \int_0^{\beta_0} d\beta' e(\beta') + \int_{\beta_0}^\beta d\beta' e(\beta') \quad (\text{H5})$$

$$= I_{\text{F}} + I_{\text{N}}, \quad (\text{H6})$$

where  $I_{\text{F}}$  and  $I_{\text{N}}$  are the integrals over the very high and the intermediate temperature regimes, respectively. We set  $\beta_0 = 1.68 \times 10^{-3}$  in this work. The integral  $I_{\text{N}}$  can be performed by usual numerical integration. To obtain  $I_{\text{F}}$  we fit the potential energy data to a polynomial function, then analytically integrate the function, which enables us to avoid the numerical integration of the diverging  $e(\beta \rightarrow 0)$  [62, 63]. In a three dimensional system of particles interacting via  $v(r) \propto r^{-12}$ , the high temperature expansion of the potential energy reads:

$$e(\beta) = A\beta^{-3/4} + B\beta^{-1/2} + C\beta^{-1/4} + D + \dots, \quad (\text{H7})$$

where  $A$ ,  $B$ ,  $C$ , and  $D$  are constants. Using Eqs. (H6) and (H7), we get

$$I_{\text{F}} = 4A\beta_0^{1/4} + 2B\beta_0^{1/2} + (4/3)C\beta_0^{3/4} + D\beta_0 + \dots. \quad (\text{H8})$$

Therefore, we can compute  $I_F$  from the fitting parameters,  $A$ ,  $B$ ,  $C$ , and  $D$ .

To obtain  $A$ ,  $B$ ,  $C$ , and  $D$  by the fitting, we rewrite Eq. (H7) as

$$\beta^{3/4}e(\beta) = A + B\left(\beta^{1/4}\right) + C\left(\beta^{1/4}\right)^2 + D\left(\beta^{1/4}\right)^3 + \dots \quad (\text{H9})$$

Thus, we use a polynomial fit for the  $\beta^{3/4}e(\beta)$  vs.  $(\beta^{1/4})$  plot as shown in Fig. S8(A). We perform the quadratic function using  $A$ ,  $B$ , and  $C$ , in this model.

In Fig. S8(B), we show the result for  $s_{\text{tot}} - s_{\text{mix}}$ . The dashed line corresponds to the Rosenfeld-Tarazona (RT) expression [64] ( $e(T) \propto T^{3/5}$  or  $s_{\text{tot}} + 3 \ln \Lambda \propto T^{-2/5}$ ) which enables us to perform extrapolation toward lower temperatures. We confirm that the RT expression works very well in our simulation range as shown in Fig. S8(B).

### 3. Vibrational entropy

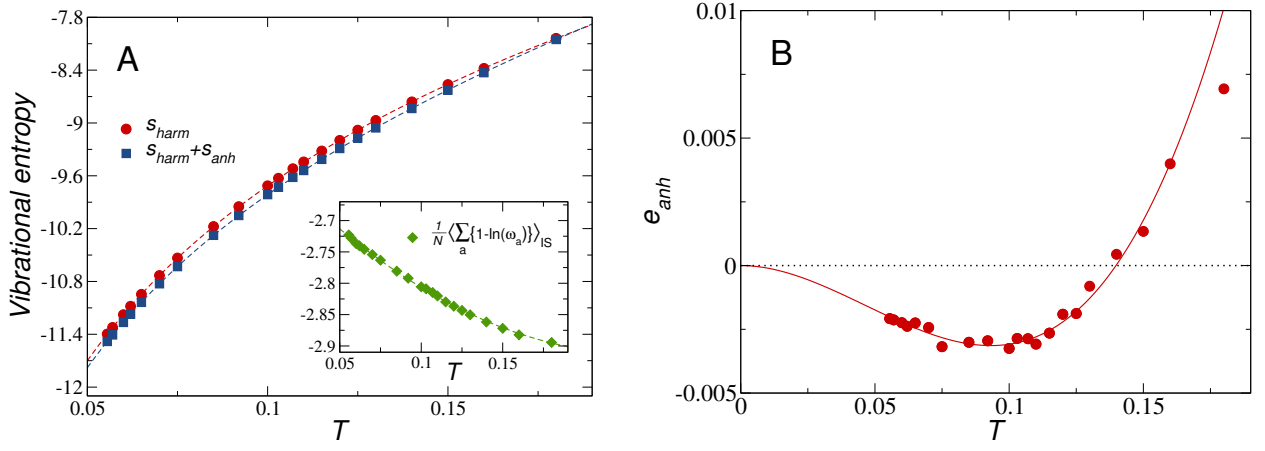


FIG. S9. (A) The harmonic vibrational entropy  $s_{\text{harm}}$  and anharmonic correction  $s_{\text{anh}}$ . The dashed curves are obtained by a quadratic fit for  $(1/N) \sum_{a=1}^{3N} \{1 - \ln(\omega_a)\}$ . (B) The anharmonic contribution of the potential energy  $e_{\text{anh}}(T)$ . The solid line is obtained by a fitting in a range  $T \in [0.0555, 0.125]$

We compute the vibrational entropy  $s_{\text{vib}}$  by

$$s_{\text{vib}} = s_{\text{harm}} + s_{\text{anh}}, \quad (\text{H10})$$

where  $s_{\text{harm}}$  and  $s_{\text{anh}}$  are the harmonic vibrational entropy and anharmonic correction, respectively [65].  $s_{\text{harm}}$  is computed [10, 49] by

$$s_{\text{harm}}(\beta) = \frac{1}{N} \left\langle \sum_{a=1}^{3N} \{1 - \ln(\beta \hbar \omega_a)\} \right\rangle_{\text{IS}}, \quad (\text{H11})$$

where  $\langle \dots \rangle_{\text{IS}}$  denotes an average over the inherent structure obtained by the conjugate gradient method, and  $\omega_a = \sqrt{\lambda_a/m}$ , and  $\lambda_a$  is the eigenvalue of the Hessian. In Fig. S9(A), we show  $s_{\text{harm}}$  as a function of  $T$ . In order to extrapolate  $s_{\text{harm}}$  toward lower temperatures, we perform the commonly used extrapolation [10, 66]; fitting  $(1/N) \langle \sum_{a=1}^{3N} \{1 - \ln(\omega_a)\} \rangle_{\text{IS}}$  to a polynomial in  $T$  of degree 2.

We evaluate the anharmonic contribution  $s_{\text{anh}}$  [65]. The anharmonic contribution of the potential energy  $e_{\text{anh}}(T)$  is given by

$$e_{\text{anh}}(T) = e(T) - e_{\text{IS}}(T) - \frac{3}{2}T, \quad (\text{H12})$$

where  $e_{\text{IS}}$  is the inherent structure energy. The last term is the energy of the harmonic vibration. Using  $e_{\text{anh}}(T)$ ,  $s_{\text{anh}}(T)$  is given by

$$s_{\text{anh}}(T) = \int_0^T dT' \frac{1}{T'} \frac{\partial e_{\text{anh}}(T')}{\partial T'}. \quad (\text{H13})$$



We assumed that  $s_{\text{anh}}(T = 0) = 0$  to derive the above equation, *i.e.*, there is no anharmonic contribution at  $T = 0$ .  $e_{\text{anh}}(T)$  can be expanded by  $T$  around zero temperature.

$$e_{\text{anh}}(T) = \sum_{k=2} a_k T^k, \quad (\text{H14})$$

where  $a_k$  is a  $T$  independent coefficient. We also assume the linear term in Eq. (H14) is zero,  $a_1 = 0$ , which means that the anharmonic contribution for the specific heat vanishes at  $T = 0$ . Substituting Eq. (H14) to Eq. (H13), we obtain

$$s_{\text{anh}}(T) = \sum_{k=2} \frac{k}{k-1} a_k T^{k-1}. \quad (\text{H15})$$

Note that the sign of the coefficients  $a_k$  determine the sign of  $s_{\text{anh}}$ .

In Fig. S9(B), we show fitting of  $e_{\text{anh}}$  using  $a_2$  and  $a_3$ . We show  $s_{\text{harm}} + s_{\text{anh}}$  in Fig. S9(A). We find that obtained anharmonic contribution is small,  $|s_{\text{anh}}| < 0.08$ , in our range of interest, showing that details of the above procedure are not very important.

#### 4. Mixing entropy

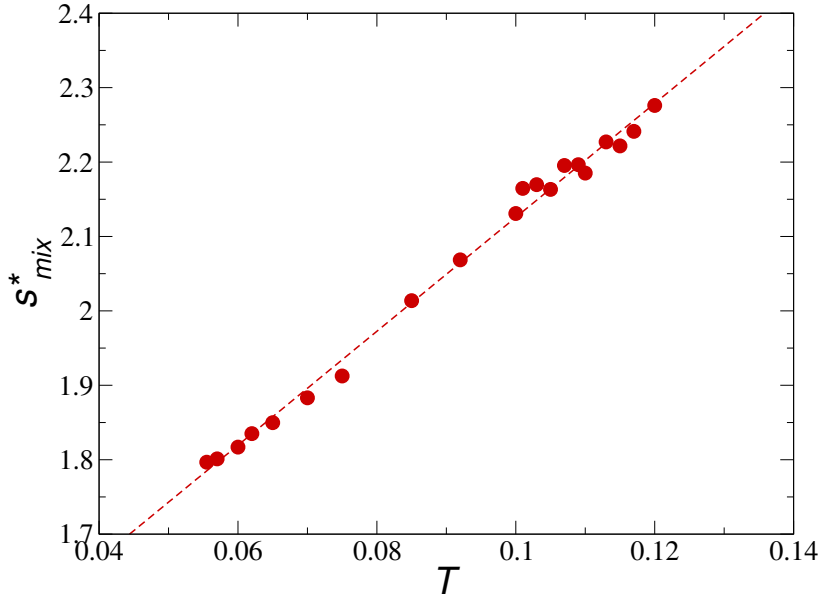


FIG. S10. The effective mixing entropy  $s_{\text{mix}}^*$ . The dashed line is an empirical fit to a straight line.

We include the effective mixing entropy  $s_{\text{mix}}^*$  obtained by an independent simulation in Ref. [25]. In contrast to the hard sphere system,  $s_{\text{mix}}^*$  of the soft sphere system has a weakly linear temperature dependence as shown in Fig. S10. Thus, we perform a linear fit and get  $s_{\text{mix}}^* = b_0 + b_1 T$ , where  $b_0 = 1.3601$  and  $b_1 = 7.6565$ .

#### 5. Configurational entropy

We show the configurational entropy  $s_{\text{conf}}$  in Fig. S11. The data demonstrates that  $s_{\text{conf}}$  decreases further below the experimental glass transition, and does not show any sign of a crossover, bending, or saturation. Also, combining the extrapolations for  $s_{\text{tot}}$ ,  $s_{\text{harm}}$ ,  $s_{\text{anh}}$ , and  $s_{\text{mix}}^*$ , we extrapolate  $s_{\text{conf}}$  down to zero, which estimates the Kauzmann transition around  $T_K \simeq 0.04$ . Therefore, we conclude from this analysis that the data analysed in the main text for a hard sphere potential are not specific to the form of the pair interaction, and confirm that our methods apply equally well to models of supercooled liquids characterized by continuous pair potentials.

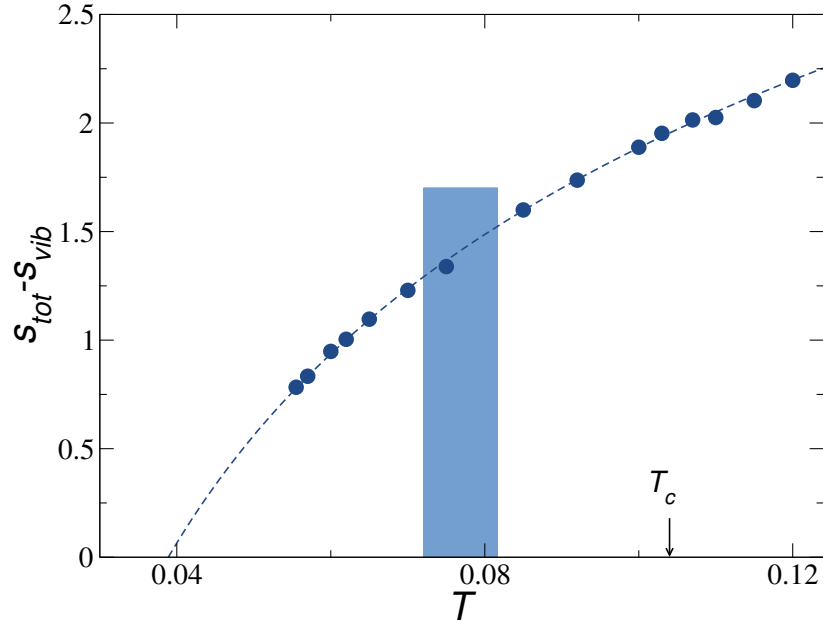


FIG. S11. The configurational entropy of the soft spheres,  $s_{\text{conf}} = s_{\text{tot}} - s_{\text{vib}}$ . The glass ceiling is indicated in the blue region and the dashed curve is an extrapolation based on fits of the individual terms. The mode-coupling crossover temperature  $T_c$  is shown by the vertical arrow.

- 
- [1] W. Kauzmann, *Chem. Rev.* **43**, 219 (1948).
  - [2] T. R. Kirkpatrick, D. Thirumalai, and P. G. Wolynes, *Phys. Rev. A* **40**, 1045 (1989).
  - [3] P. Charbonneau, J. Kurchan, G. Parisi, P. Urbani, and F. Zamponi, *Nat. Commun.* **5**, 3725 (2014).
  - [4] L. Berthier and G. Biroli, *Rev. Mod. Phys.* **83**, 587 (2011).
  - [5] F. H. Stillinger, *J. Chem. Phys.* **88**, 7818 (1988).
  - [6] J. P. Garrahan and D. Chandler, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 9710 (2003).
  - [7] J. Dyre, *Rev. Mod. Phys.* **78**, 953 (2006).
  - [8] M. Wyart, *Phys. Rev. Lett.* **104**, 095901 (2010).
  - [9] P. Rabochiy, P. G. Wolynes, and V. Lubchenko, *J. Phys. Chem. B* **117**, 15204 (2013).
  - [10] F. Sciortino, W. Kob, and P. Tartaglia, *Phys. Rev. Lett.* **83**, 3214 (1999).
  - [11] L. Berthier and D. Coslovich, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 11668 (2014).
  - [12] N. Petzold, B. Schmidtke, R. Kahlau, D. Bock, R. Meier, B. Micko, D. Kruk, and E. Rössler, *J. Chem. Phys.* **138**, 12A510 (2013).
  - [13] G. Brambilla, D. El Masri, M. Pierno, L. Berthier, L. Cipelletti, G. Petekidis, and A. B. Schofield, *Phys. Rev. Lett.* **102**, 085703 (2009).
  - [14] J.-L. Barrat, J.-N. Roux, and J.-P. Hansen, *Chem. Phys.* **149**, 197 (1990).
  - [15] W. Kob and H. C. Andersen, *Phys. Rev. Lett.* **73**, 1376 (1994).
  - [16] L.-M. Martinez and C. Angell, *Nature* **410**, 663 (2001).
  - [17] T. S. Grigera and G. Parisi, *Phys. Rev. E* **63**, 045102 (2001).
  - [18] L. Berthier, D. Coslovich, A. Ninarello, and M. Ozawa, *Phys. Rev. Lett.* **116**, 238002 (2016).
  - [19] L. Angelani and G. Foffi, *J. Phys. Cond. Matter* **19**, 256207 (2007).
  - [20] L. Berthier, P. Charbonneau, and S. Yaida, *J. Chem. Phys.* **144**, 024501 (2016).
  - [21] P. Pusey and W. Van Megen, *Nature* **320**, 340 (1986).
  - [22] L. Berthier and T. A. Witten, *Phys. Rev. E* **80**, 021502 (2009).
  - [23] D. Frenkel and A. J. Ladd, *J. Chem. Phys.* **81**, 3188 (1984).
  - [24] D. Frenkel, *Mol. Phys.* **112**, 2325 (2014).
  - [25] M. Ozawa and L. Berthier, *J. Chem. Phys.* **146**, 014502 (2017).
  - [26] S. Franz and G. Parisi, *Phys. Rev. Lett.* **79**, 2486 (1997).
  - [27] G. Adam and J. H. Gibbs, *J. Chem. Phys.* **43**, 139 (1965).
  - [28] J.-P. Bouchaud and G. Biroli, *J. Chem. Phys.* **121**, 7347 (2004).
  - [29] G. Biroli, J.-P. Bouchaud, A. Cavagna, T. S. Grigera, and P. Verrocchio, *Nat. Phys.* **4**, 771 (2008).
  - [30] G. Biroli and R. Monasson, *Europhys. Lett.* **50**, 155 (2000).
  - [31] S. Albert, T. Bauer, M. Michl, G. Biroli, J.-P. Bouchaud, A. Loidl, P. Lunkenheimer, R. Tourbot, C. Wiertel-Gasquet, and F. Ladieu, *Science* **352**, 1308 (2016).
  - [32] M. P. Allen and D. J. Tildesley, *Computer simulation of liquids* (Oxford University Press, 1989).
  - [33] A. Santos, S. B. Yuste, and M. L. de Haro, *J. Chem. Phys.* **123**, 234512 (2005).
  - [34] D. Gazzillo and G. Pastore, *Chem. Phys. Lett.* **159**, 388 (1989).
  - [35] P. Sindzingre, C. Massobrio, G. Ciccotti, and D. Frenkel, *Chem. Phys.* **129**, 213 (1989).

- [36] L. Santen and W. Krauth, arXiv preprint cond-mat/0107459 (2001).
- [37] S. Pronk and D. Frenkel, Phys. Rev. E **69**, 066123 (2004).
- [38] D. Frenkel and B. Smit, *Understanding Molecular Simulation* (Academic Press, New York, ed. 2., 2001).
- [39] Y. Brumer and D. R. Reichman, J. Phys. Chem. B **108**, 6832 (2004).
- [40] R. Gutiérrez, S. Karmakar, Y. G. Pollack, and I. Procaccia, Europhys. Lett. **111**, 56009 (2015).
- [41] S. Sastry, P. G. Debenedetti, and F. H. Stillinger, Nature **393**, 554 (1998).
- [42] P. Charbonneau, Y. Jin, G. Parisi, and F. Zamponi, Proc. Natl. Acad. Sci. U.S.A. **111**, 15025 (2014).
- [43] T. Boublik, J. Chem. Phys. **53**, 471 (1970).
- [44] G. A. Mansoori, N. F. Carnahan, K. E. Starling, and T. W. Leland Jr, J. Chem. Phys. **54**, 1523 (1971).
- [45] L. Berthier, P. Charbonneau, Y. Jin, G. Parisi, B. Seoane, and F. Zamponi, Proc. Natl. Acad. Sci. U.S.A. **113**, 8397 (2016).
- [46] Y. S. Elmatad, D. Chandler, and J. P. Garrahan, J. Phys. Chem. B **113**, 5563 (2009).
- [47] M. Isobe, A. S. Keys, D. Chandler, and J. P. Garrahan, Phys. Rev. Lett. **117**, 145701 (2016).
- [48] R. Richert and C. Angell, J. Chem. Phys. **108**, 9016 (1998).
- [49] S. Sastry, Nature **409**, 164 (2001).
- [50] J. Salacuse and G. Stell, J. Chem. Phys. **77**, 3714 (1982).
- [51] P. Sollich, J. Phys. Cond. Matter **14**, R79 (2001).
- [52] F. H. Stillinger and T. A. Weber, J. Chem. Phys. **83**, 4767 (1985).
- [53] N. Xu, J. Blawdziewicz, and C. S. O'Hern, Phys. Rev. E **71**, 061306 (2005).
- [54] K. W. Desmond and E. R. Weeks, Phys. Rev. E **80**, 051305 (2009).
- [55] L. Berthier, Phys. Rev. E **88**, 022313 (2013).
- [56] L. Berthier and R. L. Jack, Phys. Rev. Lett. **114**, 205701 (2015).
- [57] K. Hukushima and K. Nemoto, J. Phys. Soc. Jpn **65**, 1604 (1996).
- [58] H. Fukunishi, O. Watanabe, and S. Takada, J. Chem. Phys. **116**, 9058 (2002).
- [59] H. G. Katzgraber, S. Trebst, D. A. Huse, and M. Troyer, J. Stat. Mech. **2006**, P03018 (2006).
- [60] T. Vogel and D. Perez, Phys. Rev. Lett. **115**, 190602 (2015).
- [61] A. Cavagna, T. S. Grigera, and P. Verrocchio, J. Chem. Phys. **136**, 204502 (2012).
- [62] B. Coluzzi, G. Parisi, and P. Verrocchio, J. Chem. Phys. **112**, 2933 (2000).
- [63] M. Ozawa, W. Kob, A. Ikeda, and K. Miyazaki, Proc. Natl. Acad. Sci. U.S.A. **112**, 6914 (2015).
- [64] Y. Rosenfeld and P. Tarazona, Mol. Phys. **95**, 141 (1998).
- [65] F. Sciortino, J. Stat. Mech. **2005**, P05015 (2005).
- [66] E. Flenner and G. Szamel, Phys. Rev. E **73**, 061505 (2006).